

QY 121 HOETPVYLGATAGMRLLRMESEELADRVLDYVERSLSNYPEDFOGARIIITGOEGAYGMI 180
DB 181 TINYLGKFSQKTRMFSTIYPETENNQETFGALDLGASQVTFVPONOTIESPDNALQFR 240
QY 181 TINYLGKFSQKTRMFSTIYPETENNQETFGALDLGASQVTFVPONOTIESPDNALQFR 240
DB 241 LYKGDYNYTHSHFLCYGKDQALMOKLAKDIOVASNEILRDPCHFGYKRVVNSDLYKTP 300
QY 241 LYKGDYNYTHSHFLCYGKDQALMOKLAKDIOVASNEILRDPCHFGYKRVVNSDLYKTP 300
DB 301 CTKEFEMTLPEQOFEIQTGIGNYQOCHQSTILELFTNTSYCYSCAFNGIFLPLQGDGFCAF 360
QY 301 CTKEFEMTLPEQOFEIQTGIGNYQOCHQSTILELFTNTSYCYSCAFNGIFLPLQGDGFCAF 360
DB 361 SAFYVFMKFLNTSEKVSQEKYTEMMKRCQAPWEIKTSYAGYKEXLSYCFSGTYIL 420
QY 361 SAFYVFMKFLNTSEKVSQEKYTEMMKRCQAPWEIKTSYAGYKEXLSYCFSGTYIL 420
DB 421 SLLQGYHFTADSWEHIFRIGKIQSDAGWTGLGYMLNTNMIPAEQPLSTPLSHSTYVFL 480
QY 421 SLLQGYHFTADSWEHIFRIGKIQSDAGWTGLGYMLNTNMIPAEQPLSTPLSHSTYVFL 480
DB 481 MYLFSVLFTVAIIIGLLIFHKPSYFKDMV 510
QY 481 MYLFSVLFTVAIIIGLLIFHKPSYFKDMV 510

RESULT 2

ID W04264 standard; Protein: 510 AA.
AC W04264;
DT 29-MAY-1997 (first entry)
DE Human CD39 protein.
KW Human: lymphocyte activation marker; gene therapy;
KW ATP diphosphohydrolase; transgenic animal; pig; porcine; graft;
KW donor; transplantation; endothelial cell; prosthetic device;
KW platelet aggregation; inhibition; intravascular.
OS Homo sapiens.
PN W09630532-A1.
PD 03-OCT-1996.
PE 22-MAR-1996; E01270.
PR 24-MAR-1995; US-410371.
PR 12-FEB-1996; US-600383.
PA (NEMO-) NEW ENGLAND DEACONESS HOSPITAL.
PA (SANO-) SANDOZ LTD.
PI Bach FH, Robson S;
DR WPI: 96-455377/45.
DR N-PSDB: T33966.
PT Gene therapy of inflammatory or immunological stimulation of
PT platelet aggregation - using CD39 protein with ATP
PT d1:phospho:hydrolase activity; useful for preventing or alleviating
PT thrombotic condition in mammalian subject
PS Claim 3: Page 39: 65pp: English.
CC Non-human transgenic or somatic recombinant mammals, whose cells
CC contain a heterologous DNA encoding a polypeptide (especially human
CC CD39 protein) having ATP-diphosphohydrolase activity under cellular
CC activating conditions is claimed. In particular the animal is a pig
CC and its cells (or tissues or organs) can be used for transplantation.
CC DNA coding for human CD39 is also useful for genetically modifying a
CC mammalian cell to render it less susceptible to an inflammatory or
CC immunological stimulus and platelet aggregation. The modified cells
CC can be used to prevent or alleviate a thrombotic condition.
CC The present sequence is that of the human CD39 protein and was
CC disclosed in J Immunol. 153 (8) (1994) 3574-3584.
SQ Sequence 510 AA;

Query Match 100.0%; Score 3806; DB 1; Length 510;
Best local similarity 100.0%; Pred. No. 0.00e+00;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 MEDRESNWKTFCKSNLIIATIGFSSIIIVIALAVGLTQNKALPENVKYGIYLDGSSHT 60
QY 1 MEDRESNWKTFCKSNLIIATIGFSSIIIVIALAVGLTQNKALPENVKYGIYLDGSSHT 60

DB 61 SLTYIKPAPEKENDTGVAHQVEECRYKGPISKFVQKVNIEIGIYLTDCMERAREVIPSQ 120
QY 61 SLTYIKPAPEKENDTGVAHQVEECRYKGPISKFVQKVNIEIGIYLTDCMERAREVIPSQ 120
DB 121 HOETPVYLGATAGMRLLRMESEELADRVLDYVERSLSNYPEDFOGARIIITGOEGAYGMI 180
QY 121 HOETPVYLGATAGMRLLRMESEELADRVLDYVERSLSNYPEDFOGARIIITGOEGAYGMI 180
DB 181 TINYLGKFSQKTRMFSTIYPETENNQETFGALDLGASQVTFVPONOTIESPDNALQFR 240
QY 181 TINYLGKFSQKTRMFSTIYPETENNQETFGALDLGASQVTFVPONOTIESPDNALQFR 240
DB 241 LYKGDYNYTHSHFLCYGKDQALMOKLAKDIOVASNEILRDPCHFGYKRVVNSDLYKTP 300
QY 241 LYKGDYNYTHSHFLCYGKDQALMOKLAKDIOVASNEILRDPCHFGYKRVVNSDLYKTP 300
DB 301 CTKEFEMTLPEQOFEIQTGIGNYQOCHQSTILELFTNTSYCYSCAFNGIFLPLQGDGFCAF 360
QY 301 CTKEFEMTLPEQOFEIQTGIGNYQOCHQSTILELFTNTSYCYSCAFNGIFLPLQGDGFCAF 360
DB 361 SAFYVFMKFLNTSEKVSQEKYTEMMKRCQAPWEIKTSYAGYKEXLSYCFSGTYIL 420
QY 361 SAFYVFMKFLNTSEKVSQEKYTEMMKRCQAPWEIKTSYAGYKEXLSYCFSGTYIL 420
DB 421 SLLQGYHFTADSWEHIFRIGKIQSDAGWTGLGYMLNTNMIPAEQPLSTPLSHSTYVFL 480
QY 421 SLLQGYHFTADSWEHIFRIGKIQSDAGWTGLGYMLNTNMIPAEQPLSTPLSHSTYVFL 480
DB 481 MYLFSVLFTVAIIIGLLIFHKPSYFKDMV 510
QY 481 MYLFSVLFTVAIIIGLLIFHKPSYFKDMV 510

RESULT 3

ID W67846 standard; Protein: 323 AA.
AC W67846;
DT 25-MAR-1999 (first entry)
DE Human secreted protein encoded by gene 40 clone HTOEU03.
KW Human: secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; osteoclast; thymus;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
OS Homo sapiens.
PN W09842738-A1.
PD 01-OCT-1998.
PE 19-MAR-1998; U05311.
PR 30-MAY-1997; US-050937.
PR 21-MAR-1997; US-041276.
PR 21-MAR-1997; US-041277.
PR 21-MAR-1997; US-041281.
PR 21-MAR-1997; US-042344.
PR 30-MAY-1997; US-048069.
PR 30-MAY-1997; US-048094.
PR 30-MAY-1997; US-048095.
PR 30-MAY-1997; US-048096.
PR 30-MAY-1997; US-048099.
PR 30-MAY-1997; US-048131.
PR 30-MAY-1997; US-048135.
PR 30-MAY-1997; US-048154.
PR 30-MAY-1997; US-048160.
PR 30-MAY-1997; US-048186.
PR 30-MAY-1997; US-048187.
PR 30-MAY-1997; US-048188.
PR 30-MAY-1997; US-048350.
PR 30-MAY-1997; US-048351.
PR 30-MAY-1997; US-048352.
PR 30-MAY-1997; US-048355.
PR 05-AUG-1997; US-054804.
PA (HUMA-) HUMAN GENOME SCI INC.


```

FT      misc_difference 243 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 269 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 285 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 312 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 319 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 344 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 358 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 378 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 386 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 465 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 473 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 473 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 494 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 555 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 593 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 602 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 609 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 621 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 724 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 736 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 739 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 786 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 841 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 924 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 934 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 1017 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 1034 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 1127 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 1147 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 1178 /note= "corresponds to stop codon in DNA sequence"
FT      misc_difference 1178 /note= "corresponds to stop codon in DNA sequence"
PN      WC0632476-A1.
PD      17-OCT-1996.
PR      12-APR-1996; CA0214.
PR      13-APR-1995; US-421701.
PA      (MOUN ) MOUNT SINAI HOSPITAL CORP.
PI      Labes M, Lozano A, Roach A, Roder J;
DR      WPI; 96-477127/47.
DR      N-PSDB: T38484.
PT      Assay for substance that modulates response of neuronal cells - and
PT      neurite growth associated protein, Petrin, useful in conditions
PT      involving nerve damage resulting from traumatic injury, stroke or
PT      CNS degenerative disorders

```

```

PS      Claim 9; Page 57-61; 119pp; English.
CC      Rat petrin (W04336) is a protein involved in modulating neurite
CC      growth inhibition. The amino sequence was deduced from a cDNA
CC      clone (T38484) derived from an adult rat brain cDNA library; no
CC      coding sequence was indicated. Petrin is a new member of the
CC      protein phosphatase 2C family, and is expressed in neurons in brain
CC      tissue, partic. in the Purkinje cells of the cerebellum. Petrin,
CC      and antibodies raised against it, can be used to modulate neurite
CC      growth and axonal regeneration.
SQ      Sequence 1196 AA;

Query Match      2.8%; Score 106; DB 1; Length 1196;
Best Local Similarity 31.4%; Pred. No. 1.51e+01;
Matches 16; Conservative 15; Mismatches 18; Indels 2; Gaps 2;

DB      503 KYGQSHCVFOLFQSGFPLPISCPLEFSIFFCGCGVQSGFPGFVL-FL 552
QY      321 NYQDCHQSTLELFTNSYCPYSCAFN-GFLPLQGDGFAFSAEFYWKFL 370

RESULT      7
ID      W04339 standard; Protein; 20 AA.
AC      W04339;
DT      29-DEC-1996 (first entry)
DE      ATP di phosphohydrolase fragment from pig pancreas.
KW      ATP di phosphohydrolase; ATPase; bovine aorta; pig pancreas;
KW      apyrase; CD39; lymphoid cell activation antigen; enzyme;
KW      platelet aggregation; thrombogenicity; anti-haemostatic.
OS      Sus scrofa.
PN      W09632471-A2.
PD      17-OCT-1996.
PR      10-APR-1996; CA0223.
PR      10-APR-1995; US-419204.
PA      (UYSH ) UNIV SHERBROOKE.
PI      Beaudoin AR, Sevigny J;
DR      WPI; 96-477122/47.
PT      Isolated ATP di:phospho:hydrolase enzymes - have anti-haemostatic
PT      activity, useful for reducing platelet aggregation and
PT      thrombogenicity
PS      Claim 3; Page 46; 60pp; English.
CC      The bovine aorta (W04335 to W04338 and W04340) and porcine
CC      pancreatic (W04339) ATPases have been partially sequenced.
CC      The sequences have been found to be highly homologous to a human
CC      lymphoid cell activation antigen designated CD39 (Maliszewski et
CC      al. (1994). J Immunol. : 3574-3583). The complete sequences of the
CC      ATPases types I and II have not been obtained yet. Assuming that
CC      the CD39 gene product is an ATPase type II, the use of CD39
CC      in the reduction of platelet aggregation and of thrombogenicity
CC      may be contemplated, as well as a process of making ATPases
CC      using the CD39 sequence (T38516).
SQ      Sequence 20 AA;

Query Match      2.7%; Score 104; DB 1; Length 20;
Best Local Similarity 78.9%; Pred. No. 2.06e+01;
Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

DB      2 SPTQETGALDYGASTOV 20
QY      203 TNNQETFGALDYGASTOV 221

RESULT      8
ID      W20877 standard; Protein; 95 AA.
AC      W20877;
DT      18-JUL-1997 (first entry)
DE      H. pylori cytoplasmic protein, 13ee10215orf82.
KW      Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
KW      identification; binding compound; bacterium; life cycle; activator;
KW      bacteria; inhibitor; duodenal ulcer disease; chronic gastritis;
KW      diagnosis.
OS      Helicobacter pylori.
PN      W09640893-A1.
PD      19-DEC-1996.

```

PF 06-JUN-1996; U09122.
PR 07-JUN-1995; US-487032.
PR 01-APR-1996; US-630405.
PA (ASTR) ASTRA AB.
PI Berglindh O, Smith D, Møllgaard BL;
DR WPI; 97-052306/05.
DR N-PSDB; T68130.
PT Helicobacter pylori nucleic acid sequences and related
PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
PT Infection, and to detect Helicobacter
PS Claim 61; Page 1278-79; 1481pp; English.
CC This sequence represents a H. pylori cytoplasmic protein.
CC The protein may be used in a vaccine to prevent or treat H. pylori
CC infection or to identify H. pylori polypeptide binding compounds,
CC useful as potential H. pylori life cycle activators or inhibitors.
CC The genomic sequence of H. pylori (ATCC 55679) was determined from
CC overlapping contigs generated by mechanically shearing the bacterial
CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,
CC and the predicted coding regions defined by computer evaluation. To
CC identify likely H. pylori antigens for vaccine development, the amino
CC acid sequences predicted from various ORF were analysed for significant
CC homology to other known or exported membrane proteins. Having identified
CC and determined the sequences of interest, particular regions can be
CC isolated from H. pylori by PCR amplification for recombinant polypeptide
CC production, e.g. in E. coli hosts.
SQ Sequence 95 AA;

Query Match 2.7%; Score 103; DB 1; Length 95;
Best Local Similarity 34.0%; Pred. No. 2.41e+01;
Matches 18; Conservative 16; Mismatches 14; Indels 5; Gaps 5;

Db 24 GYM-VSLUGALKRTPCRREFLKALFAIF-YHAVNNPLTQCPRPHQREFSS 74
QY 286 GYKRVNVSDLYKTPCKRFFEM-TLPPQGFQIGIGNY-QQCH-QSILELFNT 335

RESULT 9
ID R85880 standard; Protein; 439 AA.
AC R85880;
DT 13-SEP-1996 (first entry)
DE WD-40 domain-contg. YC07 protein.
KW WD40 repeat region; beta-transducin; protein-protein interaction; drug;
KW intracellular signalling; protein kinase C; homology; motif; modulator;
KW receptors of activated protein kinase; enzyme activity; isozyme; human.
OS Synthetic.
PN WO9521252-A2.
PD 10-AUG-1995.
PF 31-JAN-1995; U01210.
PR 01-FEB-1994; US-190802.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Mochly-Rosen D, Ron D;
DR WPI; 95-283772/37.
PT New WD-40 (beta-transducin)-derived polypeptide(s) - which alter the
PT activity of a protein, eg. protein kinase C, which interacts with a
PT protein contg. a WD-40 region.
PS Example 5; Page 166-168; 351pp; English.
CC Proteins R85831-92 are protein which contain at least one WD-40 (also
CC called beta-transducin homologous) amino acid repeat motifs. The WD-40
CC regions are involved in protein-protein interactions between proteins
CC involved in intracellular signalling. An example of such an interaction
CC is between protein kinase C and receptors of activated protein kinase
CC (RACK), esp. RACK-1 (R85850). Proteins R85851-82 were isolated based on
CC homology with beta-transducin, whereas proteins R85882-92 were isolated
CC based on homology with the WD-40 consensus sequence (R85893). The
CC proteins were used to construct the peptides R84928-R85063 and
CC R85786-R85842. The peptides can be used to identify target proteins of
CC contg. WD-40 motifs, as modulators of enzyme esp. isozyme, activity of
CC proteins involved in protein-protein interaction and to screen for drugs
CC that will affect protein-protein interaction involving WD-40 domains.
SQ Sequence 439 AA;

Query Match 2.7%; Score 102; DB 1; Length 439;
Best Local Similarity 26.7%; Pred. No. 2.81e+01;

Matches 16; Conservative 16; Mismatches 25; Indels 3; Gaps 3;

Db 180 AYPDGRVTVASDEGKTKWDITSGFLATFEHTSSVTVAVQAKQVWFSS-SLDGT 238
QY 160 PFDFGARIITGOEGAV-GW-ITINVLGKFSQKTRFSTVPEYTNQETFGALDLOGA 217

RESULT 10
ID W55029 standard; Protein; 481 AA.
AC W55029;
DT 17-SEP-1998 (first entry)
DE G-protein coupled receptor, short form.
KW G-protein coupled receptor; gene therapy; abnormality detection;
KW short form; human.
OS Homo sapiens.
PN EP-845529-A2.
PD 03-JUN-1998.
PF 27-OCT-1997; 308562.
PR 29-OCT-1996; JP-286823.
PA (TAKE) TAKEDA CHEM IND LTD.
PI Fukusumi S, Hinuma S, Kawamata Y;
DR WPI; 98-288746/26.
DR N-PSDB; V27174.
PT New human G-protein coupled receptor protein - and corresponding
PT DNA, ligands, antibodies, etc
PS Claim 1; Page 42-44; 65pp; English.
CC This sequence represents a human G-protein coupled receptor of the
CC invention. The protein or cells expressing the DNA encoding it can be
CC used to screen for agonists or antagonists of the receptor, which can be
CC used as drugs for treating various diseases (none disclosed). The DNA can
CC also be used for practice drug design based on comparisons with
CC structurally analogous ligands and receptors. DNA encoding the protein
CC can be used for gene therapy for diseases caused by a deficiency of the
CC receptor. The DNA can also be used to detect abnormalities in the gene
CC encoding the receptor. The protein or fragment can be used to determine
CC levels of receptor ligands in vivo. The antibody can be used in assays to
CC detect the protein.
SQ Sequence 481 AA;

Query Match 2.7%; Score 101; DB 1; Length 481;
Best Local Similarity 35.1%; Pred. No. 3.28e+01;
Matches 13; Conservative 14; Mismatches 8; Indels 2; Gaps 2;

Db 114 LQIONPLX-PYTESSTAYAMALALVFNAGIVGNL 149
QY 462 IPAEQPLSTPLSHSTY-VFLMWFLSVLFTVAIIQL 497

RESULT 11
ID W55030 standard; Protein; 542 AA.
AC W55030;
DT 17-SEP-1998 (first entry)
DE G-protein coupled receptor, long form.
KW G-protein coupled receptor; gene therapy; abnormality detection;
KW long form; human.
OS Homo sapiens.
PN EP-845529-A2.
PD 03-JUN-1998.
PF 27-OCT-1997; 308562.
PR 29-OCT-1996; JP-286823.
PA (TAKE) TAKEDA CHEM IND LTD.
PI Fukusumi S, Hinuma S, Kawamata Y;
DR WPI; 98-288746/26.
DR N-PSDB; V27175.
PT New human G-protein coupled receptor protein - and corresponding
PT DNA, ligands, antibodies, etc
PS Claim 2; Page 45-47; 65pp; English.
CC This sequence represents a human G-protein coupled receptor of the
CC invention. The protein or cells expressing the DNA encoding it can be
CC used to screen for agonists or antagonists of the receptor, which can be
CC used as drugs for treating various diseases (none disclosed). The DNA can
CC also be used for practice drug design based on comparisons with
CC structurally analogous ligands and receptors. DNA encoding the protein

CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Frommer, William S.
CC REGISTRATION NUMBER: 25,506
CC REFERENCE/DOCKET NUMBER: 45310-2540
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 840-3333
CC TELEFAX: (212) 840-0712
CC TELEX: 425066 CURTMS
CC INFORMATION FOR SEQ ID NO: 12:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 459 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC FRAGMENT TYPE: internal
CC SEQUENCE 459 AA; 52308 MW; 1156694 CN;
SO
Query Match 2.5%; Score 96; DB 1; Length 459;
Best Local Similarity 32.6%; Pred. No. 3.50e+01;
Matches 14; Conservative 12; Mismatches 16; Indels 1; Gaps 1;
Db 344 CYSRENNKLVKQIPINGDDIISGCLNHPGLVNIQNKIDI 386
QY 255 CYGDOALMOKLAKIQVANSNELLRDPCE-HPGYKKVNVSDL 296
RESULT 7
ID US-08-923-772-2 STANDARD; PRT; 523 AA.
XX xxxxxx
AC
DE Sequence 2, Application US/08923772
XX
CC Sequence 2, Application US/08923772
CC Patent No. 5972651
CC GENERAL INFORMATION:
CC APPLICANT: Black, Michael T.
CC TITLE OF INVENTION: NOVEL FIB
CC NUMBER OF SEQUENCES: 2
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Dechert Price & Rhoads
CC STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CC CITY: Philadelphia
CC STATE: PA
CC COUNTRY: US
CC ZIP: 19103
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC OPERATING SYSTEM: DOS
CC SOFTWARE: FastSeq for Windows Version 2.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/923,772
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Dickinson, Todd O
CC REGISTRATION NUMBER: 28,354
CC REFERENCE/DOCKET NUMBER: G410080
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 215-994-2252
CC TELEFAX: 215-994-2222
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 523 amino acids
CC

CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC SEQUENCE 523 AA; 57775 MW; 1280224 CN;
SO
Query Match 2.5%; Score 94; DB 2; Length 523;
Best Local Similarity 20.7%; Pred. No. 4.72e+01;
Matches 29; Conservative 40; Mismatches 64; Indels 7; Gaps 7;
Db 196 LQYDELLMNEIDVKNVLAQPNELLVYDAMIGQEAANVARFENAOLEY-TGVIIITKIDGD 254
QY 27 IAVIALAVGLTONKALPENVKYGIYLDGSSHTSLYIK-WPAKENDGVHVEECR 85
Db 255 TRG-GAALSVRHITGKPIKFTGTGKRTDIEFHHDDPMSSRI-LGMDMLTLIERASQOEY 312
QY 86 VKGPGISKRFVQKYNELGIYLTDOMERAREVIP-RSQHDETPIYLGATAGMRLIRRESEEL 144
Db 313 DEOKALEMAEK-WRENTDFE 331
QY 145 AD-RVIDVERSLSNYPDFE 163
RESULT 8
ID US-08-466-589-6 STANDARD; PRT; 627 AA.
XX xxxxxx
AC
DE Sequence 6, Application US/08466589
XX
CC Sequence 6, Application US/08466589
CC Patent No. 5837489
CC GENERAL INFORMATION:
CC APPLICANT: Elliot, Kathryn J.
CC APPLICANT: Ellis, Steven B.
CC APPLICANT: Harpold, Michael M.
CC TITLE OF INVENTION: HUMAN NEURONAL NICOTINIC ACETYLCHOLINE
CC TITLE OF INVENTION: RECEPTOR COMPOSITIONS AND METHODS EMPLOYING SAME
CC NUMBER OF SEQUENCES: 12
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Brown, Martin, Haller & Mcclain
CC STREET: 1660 Union Street
CC CITY: San Diego
CC STATE: CA
CC COUNTRY: US
CC ZIP: 92101-2926
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC OPERATING SYSTEM: DOS
CC SOFTWARE: FastSeq Version 1.5
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/466,589
CC FILING DATE: June 5, 1995
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/028,031
CC FILING DATE: March 8, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Seldman, Stephanie L
CC REGISTRATION NUMBER: 33,779
CC REFERENCE/DOCKET NUMBER: 6362-9950
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 619-238-0999
CC TELEFAX: 619-238-0062
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 6:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 627 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC

CC	SEQUENCE	627 AA; 69767 MW; 2152270 CN;	MOLECULE TYPE: protein
QY	Query Match	2.5%; Score 97; DB 2; Length 627;	
CC	Best Local Similarity	26.9%; Pred. No. 3.01e+01;	
CC	Matches	14; Conservative	22; Mismatches 12; Indels 4; Gaps 4;
Db	287 GFLILTEIIPPT-SLVYPLI-GEYLFTMIEVYLSI-AIVFVLNVHHRSP 335		
QY	453 GYMLNLTNMPADPPLSTPLSHSTYV-FLMVFLSLVFLVATIGLILFHKPS 503		
CC	Sequence 6, Application US/08700636		
CC	Sequence 6, Application US/08700636		
CC	Patent No. 5910582		
CC	GENERAL INFORMATION:		
CC	APPLICANT: Elliot, Kathryn J.		
CC	APPLICANT: Ellis, Steven B.		
CC	APPLICANT: Harpold, Michael M.		
CC	TITLE OF INVENTION: HUMAN NEURONAL NICOTINIC ACETYLCHOLINE		
CC	TITLE OF INVENTION: RECEPTOR COMPOSITIONS AND METHODS EMPLOYING SAME		
CC	NUMBER OF SEQUENCES: 12		
CC	CORRESPONDENCE ADDRESS:		
CC	ADDRESS: Pretty, Schroeder, Brueggemann & Clark		
CC	STREET: 444 South Flower Street, Suite 2000		
CC	CITY: Los Angeles		
CC	STATE: CA		
CC	COUNTRY: USA		
CC	ZIP: 90071		
CC	COMPUTER READABLE FORM:		
CC	MEDIUM TYPE: Floppy disk		
CC	COMPUTER: IBM PC compatible		
CC	OPERATING SYSTEM: PC-DOS/MS-DOS		
CC	SOFTWARE: PatentIn Release #1.0, Version #1.25		
CC	CURRENT APPLICATION DATA:		
CC	APPLICATION NUMBER: US/08/700,636		
CC	FILING DATE: 16-JUL-1996		
CC	CLASSIFICATION: 356		
CC	PRIOR APPLICATION DATA:		
CC	APPLICATION NUMBER: US 08/028,031		
CC	FILING DATE: 08-MAR-1993		
CC	ATTORNEY/AGENT INFORMATION:		
CC	NAME: Reiter, Stephen E.		
CC	REGISTRATION NUMBER: 31,192		
CC	REFERENCE/DOCKET NUMBER: P41 9368		
CC	TELECOMMUNICATION INFORMATION:		
CC	TELEPHONE: 619-546-4737		
CC	TELEFAX: 619-546-9392		
CC	INFORMATION FOR SEQ ID NO: 6:		
CC	SEQUENCE CHARACTERISTICS:		
CC	LENGTH: 627 amino acids		
CC	TYPE: amino acid		
CC	STRANDEDNESS: single		
CC	TOPOLOGY: unknown		
CC	MOLECULE TYPE: protein		
CC	SEQUENCE 627 AA; 69767 MW; 2152270 CN;		
Db	287 GFLILTEIIPPT-SLVYPLI-GEYLFTMIEVYLSI-AIVFVLNVHHRSP 335		
QY	453 GYMLNLTNMPADPPLSTPLSHSTYV-FLMVFLSLVFLVATIGLILFHKPS 503		
CC	Sequence 6, Application US/08700636		
CC	Sequence 6, Application US/08700636		
CC	Patent No. 5910582		
CC	GENERAL INFORMATION:		
CC	APPLICANT: Elliot, Kathryn J.		
CC	APPLICANT: Ellis, Steven B.		
CC	APPLICANT: Harpold, Michael M.		
CC	TITLE OF INVENTION: HUMAN NEURONAL NICOTINIC ACETYLCHOLINE		
CC	TITLE OF INVENTION: RECEPTOR COMPOSITIONS AND METHODS EMPLOYING SAME		
CC	NUMBER OF SEQUENCES: 12		
CC	CORRESPONDENCE ADDRESS:		
CC	ADDRESS: Pretty, Schroeder, Brueggemann & Clark		
CC	STREET: 444 South Flower Street, Suite 2000		
CC	CITY: Los Angeles		
CC	STATE: CA		
CC	COUNTRY: USA		
CC	ZIP: 90071		
CC	COMPUTER READABLE FORM:		
CC	MEDIUM TYPE: Floppy disk		
CC	COMPUTER: IBM PC compatible		
CC	OPERATING SYSTEM: PC-DOS/MS-DOS		
CC	SOFTWARE: PatentIn Release #1.0, Version #1.25		
CC	CURRENT APPLICATION DATA:		
CC	APPLICATION NUMBER: US/08/700,636		
CC	FILING DATE: 16-JUL-1996		
CC	CLASSIFICATION: 356		
CC	PRIOR APPLICATION DATA:		
CC	APPLICATION NUMBER: US 08/028,031		
CC	FILING DATE: 08-MAR-1993		
CC	ATTORNEY/AGENT INFORMATION:		
CC	NAME: Reiter, Stephen E.		
CC	REGISTRATION NUMBER: 31,192		
CC	REFERENCE/DOCKET NUMBER: P41 9368		
CC	TELECOMMUNICATION INFORMATION:		
CC	TELEPHONE: 619-546-4737		
CC	TELEFAX: 619-546-9392		
CC	INFORMATION FOR SEQ ID NO: 6:		
CC	SEQUENCE CHARACTERISTICS:		
CC	LENGTH: 627 amino acids		
CC	TYPE: amino acid		
CC	STRANDEDNESS: single		
CC	TOPOLOGY: unknown		
CC	MOLECULE TYPE: protein		
CC	SEQUENCE 627 AA; 69767 MW; 2152270 CN;		
Db	287 GFLILTEIIPPT-SLVYPLI-GEYLFTMIEVYLSI-AIVFVLNVHHRSP 335		
QY	453 GYMLNLTNMPADPPLSTPLSHSTYV-FLMVFLSLVFLVATIGLILFHKPS 503		
CC	Sequence 6, Application US/08700636		
CC	Sequence 6, Application US/08700636		
CC	Patent No. 5910582		
CC	GENERAL INFORMATION:		
CC	APPLICANT: Elliot, Kathryn J.		
CC	APPLICANT: Ellis, Steven B.		
CC	APPLICANT: Harpold, Michael M.		
CC	TITLE OF INVENTION: HUMAN NEURONAL NICOTINIC ACETYLCHOLINE		
CC	TITLE OF INVENTION: RECEPTOR COMPOSITIONS AND METHODS EMPLOYING SAME		
CC	NUMBER OF SEQUENCES: 12		
CC	CORRESPONDENCE ADDRESS:		
CC	ADDRESS: Pretty, Schroeder, Brueggemann & Clark		

RESULT	10		STANDARD;	PRT:	971 AA.
ID	US-08-480-662-2				
XX	xxxxxx				
AC					
DT					
XX					
DE	Sequence 2, Application US/08480662				
XX					
CC	Sequence 2, Application US/08480662				
CC	Patent No. 5759782				
CC	GENERAL INFORMATION:				
CC	APPLICANT: Pastan, Ira				
CC	APPLICANT: Brinkmann, Ulrich				
CC	TITLE OF INVENTION: CELLULAR APOPTOSIS SUSCEPTIBILITY PROTEIN (CSP) AND AN				
CC	NUMBER OF SEQUENCES: 10				
CC	CORRESPONDENCE ADDRESS:				
CC	ADDRESSEE: Knobbe, Martens, Olson and Bear				
CC	STREET: 620 Newport Center Drive 16th Floor				
CC	City: Newport Beach				
CC	STATE: CA				
CC	COUNTRY: USA				
CC	ZIP: 92660				
CC	COMPUTER READABLE FORM:				
CC	MEDIUM TYPE: Diskette				
CC	COMPUTER: IBM Compatible				
CC	OPERATING SYSTEM: DOS				
CC	SOFTWARE: FastSeq Version 1.5				
CC	CURRENT APPLICATION DATA:				
CC	APPLICATION NUMBER: US/08/480,662				
CC	FILING DATE: 07-JUN-1995				
CC	CLASSIFICATION: 530				
CC	PRIOR APPLICATION DATA:				
CC	APPLICATION NUMBER:				
CC	FILING DATE:				
CC	ATTORNEY/AGENT INFORMATION:				
CC	NAME: Israelson, Ned A				
CC	REGISTRATION NUMBER: 29,655				
CC	REFERENCE/DOCKET NUMBER: NIH112.001A				
CC	TELECOMMUNICATION INFORMATION:				
CC	TELEPHONE: 619-235-8550				
CC	TELEFAX: 619-235-0176				
CC	TELEX:				
CC	INFORMATION FOR SEQ ID NO: 2:				
CC	SEQUENCE CHARACTERISTICS:				
CC	LENGTH: 971 amino acids				
CC	TYPE: amino acid				
CC	STRANDEDNESS: single				
CC	TOPOLOGY: linear				
CC	MOLECULE TYPE: peptide				
CC	HYPOTHETICAL: NO				
CC	ANTI-SENSE: NO				
CC	FRAGMENT TYPE: N-terminal				
CC	ORIGINAL SOURCE:				
CC	SEQUENCE 971 AA; 110313 MW; 4888742 CN;				
SO					
	Query Match	2.5%;	Score 94;	DB 1;	Length 971;
	Best Local Similarity	28.8%;	Pred. No. 4.72e+01;		
	Matches 17;	Conservative 17;	Mismatches 21;	Indels 4;	Gaps 4;
Db	179 AFALPLTNIKFKATIGCTGHANDASA-LRILPSSLLIKSLFYSL-NFODLPEFWEGNM 235				
Oy	453 GYMLNLTNNIRAEQPLSLPSLSHTVFLNLVLFSL-LVLFTVALIIGLLIFHKPSYFKK-DM 509				
RESULT	11		STANDARD;	PRT:	971 AA.
ID	PCT-US96-09927-2				
XX	xxxxxx				
AC					
DT					
XX					

Query Match 2.4%; Score 93; DB 2; Length 175;
 Best Local Similarity 29.1%; Pred. No. 5.48e+01;
 Matches 16; Conservative 13; Mismatches 24; Indels 2; Gaps 2;
 Db 65 LNLGHQALFOEKVITLD-FSTPGRAAAVAFLSFITKDKWSEETHLSGGYLDL 118
 QY 370 LNLTFSEKVSQEKVTEEMKKFCAQPMWEIKTSYAG-VKEKYLSEYCFSGTYILSL 423

Search completed: Fri May 5 09:00:32 2000
 Job time : 81 secs.

Query Match	100.0%;	Score 3806;	DB 25;	Length 510;
Best Local Similarity	100.0%;	Pred. No. 0.00e+00;		
Matches 510;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

```
|||||
QY 1 MEDTKESNVKTFCSKNILAILGFSSIIAVTALLAVGLTONKALPENKYGVILDAGSSHT 60
Db 61 SLTYKKPAEKENDTGVVHOVEECRVKGPISKRVQKNEIGIYLTDCMEARAREVIRSQ 120
QY 61 SLTYKKPAEKENDTGVVHOVEECRVKGPISKRVQKNEIGIYLTDCMEARAREVIRSQ 120
Db 121 HOETPVYLGATAGMRLLRMESEELADRVLDVERSLSNYPEDFGARLITGOEBGAYGWI 180
QY 121 HOETPVYLGATAGMRLLRMESEELADRVLDVERSLSNYPEDFGARLITGOEBGAYGWI 180
Db 181 TINYLKFSQKTRWFSTIVPETNNOETFGALDAGASTOVTFVPQNOTIESPINALQFR 240
QY 181 TINYLKFSQKTRWFSTIVPETNNOETFGALDAGASTOVTFVPQNOTIESPINALQFR 240
Db 241 LYGDYVNYTHSFICYGKDQALMOKLAKDIOVASNEILRDCFPBGYKKVNVNSDLTKTP 300
QY 241 LYGDYVNYTHSFICYGKDQALMOKLAKDIOVASNEILRDCFPBGYKKVNVNSDLTKTP 300
Db 301 CTKRFEWTLRPPQPEIIGIGNYQOCHOSILELFTNTSYCPYSQCAFNGIFLPPLOGDGCAGF 360
QY 301 CTKRFEWTLRPPQPEIIGIGNYQOCHOSILELFTNTSYCPYSQCAFNGIFLPPLOGDGCAGF 360
Db 361 SAFYFVKFPLNLJSEKYSQEKVTEEMKKFCAQPMWEIKTSYAGVKEKYLSEYCSGTYIL 420
QY 361 SAFYFVKFPLNLJSEKYSQEKVTEEMKKFCAQPMWEIKTSYAGVKEKYLSEYCSGTYIL 420
Db 421 SLLQGYHFTADSWEHHTFGIKIGSDAGWTLGYMLNTNMIPAEOPLSPLSHSTVFL 480
QY 421 SLLQGYHFTADSWEHHTFGIKIGSDAGWTLGYMLNTNMIPAEOPLSPLSHSTVFL 480
Db 481 MVLESLVLFVTAITGLLIFHKPSYFKMDV 510
QY 481 MVLESLVLFVTAITGLLIFHKPSYFKMDV 510
```

```
RESULT 4
ID PCT-US99-22955-2 STANDARD; PRT; 510 AA.
XX AC xxxxxx
XX DT
XX DE
XX Sequence 2, Application PC/TUS9922955
XX CC
XX CC Sequence 2, Application PC/TUS9922955
XX CC GENERAL INFORMATION:
XX CC APPLICANT: Maliszewski, Charles R.
XX CC APPLICANT: Gayle III, Richard B.
XX CC APPLICANT: Price, Virginia L.
XX CC APPLICANT: Gimpel, Steven D.
XX CC APPLICANT: Immunex Corporation
XX CC TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment
XX CC FILE REFERENCE: 2879-WO
XX CC CURRENT APPLICATION NUMBER: PCT/US99/22955
XX CC EARLIER APPLICATION NUMBER: US 60/104,585
XX CC EARLIER FILING DATE: 1998-10-16
XX CC EARLIER APPLICATION NUMBER: US 60/107,466
XX CC EARLIER FILING DATE: 1998-11-06
XX CC EARLIER APPLICATION NUMBER: US 60/149,010
XX CC EARLIER FILING DATE: 1999-08-13
XX CC NUMBER OF SEQ ID NOS: 31
XX CC SOFTWARE: Patentln Ver. 2.0
XX CC SEQ ID NO 2
XX CC LENGTH: 510
XX CC TYPE: prt
XX CC ORGANISM: Homo sapiens
XX CC SEQUENCE 510 AA: 57964 MW; 1469467 CN;
SQ
```

Query Match 100.0%; Score 3806; DB 1; Length 510;
Best local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
|||||
Db 1 MEDTKESNVKTFCSKNILAILGFSSIIAVTALLAVGLTONKALPENKYGVILDAGSSHT 60
QY 1 MEDTKESNVKTFCSKNILAILGFSSIIAVTALLAVGLTONKALPENKYGVILDAGSSHT 60
Db 61 SLTYKKPAEKENDTGVVHOVEECRVKGPISKRVQKNEIGIYLTDCMEARAREVIRSQ 120
QY 61 SLTYKKPAEKENDTGVVHOVEECRVKGPISKRVQKNEIGIYLTDCMEARAREVIRSQ 120
Db 121 HOETPVYLGATAGMRLLRMESEELADRVLDVERSLSNYPEDFGARLITGOEBGAYGWI 180
QY 121 HOETPVYLGATAGMRLLRMESEELADRVLDVERSLSNYPEDFGARLITGOEBGAYGWI 180
Db 181 TINYLKFSQKTRWFSTIVPETNNOETFGALDAGASTOVTFVPQNOTIESPINALQFR 240
QY 181 TINYLKFSQKTRWFSTIVPETNNOETFGALDAGASTOVTFVPQNOTIESPINALQFR 240
Db 241 LYGDYVNYTHSFICYGKDQALMOKLAKDIOVASNEILRDCFPBGYKKVNVNSDLTKTP 300
QY 241 LYGDYVNYTHSFICYGKDQALMOKLAKDIOVASNEILRDCFPBGYKKVNVNSDLTKTP 300
Db 301 CTKRFEWTLRPPQPEIIGIGNYQOCHOSILELFTNTSYCPYSQCAFNGIFLPPLOGDGCAGF 360
QY 301 CTKRFEWTLRPPQPEIIGIGNYQOCHOSILELFTNTSYCPYSQCAFNGIFLPPLOGDGCAGF 360
Db 361 SAFYFVKFPLNLJSEKYSQEKVTEEMKKFCAQPMWEIKTSYAGVKEKYLSEYCSGTYIL 420
QY 361 SAFYFVKFPLNLJSEKYSQEKVTEEMKKFCAQPMWEIKTSYAGVKEKYLSEYCSGTYIL 420
Db 421 SLLQGYHFTADSWEHHTFGIKIGSDAGWTLGYMLNTNMIPAEOPLSPLSHSTVFL 480
QY 421 SLLQGYHFTADSWEHHTFGIKIGSDAGWTLGYMLNTNMIPAEOPLSPLSHSTVFL 480
Db 481 MVLESLVLFVTAITGLLIFHKPSYFKMDV 510
QY 481 MVLESLVLFVTAITGLLIFHKPSYFKMDV 510
```

```
RESULT 5
ID PCT-US99-22955-2 STANDARD; PRT; 464 AA.
XX AC xxxxxx
XX DT
XX DE
XX Sequence 27, Application PC/TUS9922955
XX CC
XX CC Sequence 27, Application PC/TUS9922955
XX CC GENERAL INFORMATION:
XX CC APPLICANT: Maliszewski, Charles R.
XX CC APPLICANT: Gayle III, Richard B.
XX CC APPLICANT: Price, Virginia L.
XX CC APPLICANT: Gimpel, Steven D.
XX CC APPLICANT: Immunex Corporation
XX CC TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment
XX CC FILE REFERENCE: 2879-WO
XX CC CURRENT APPLICATION NUMBER: PCT/US99/22955
XX CC EARLIER APPLICATION NUMBER: US 60/104,585
XX CC EARLIER FILING DATE: 1998-10-16
XX CC EARLIER APPLICATION NUMBER: US 60/107,466
XX CC EARLIER FILING DATE: 1998-11-06
XX CC EARLIER APPLICATION NUMBER: US 60/149,010
XX CC EARLIER FILING DATE: 1999-08-13
XX CC NUMBER OF SEQ ID NOS: 31
XX CC SOFTWARE: Patentln Ver. 2.0
XX CC SEQ ID NO 27
XX CC LENGTH: 464
XX CC TYPE: prt
XX CC ORGANISM: Artificial Sequence
XX CC FEATURE:
XX OTHER INFORMATION: Description of Artificial Sequence: Fusion
XX OTHER INFORMATION: construct of human CD39
```

Qy	98	VNEIGITVITDCMERARREVI	PPRSQHOETPVYLGATATGMLLM	MESEBELADRYLDVVERSL	157
Db	121	NYPDEOGARITITGOEGAGWITITNLL	LGKFSQKTRMF <td>SIVPYETNNQETFGALD</td> <td>180</td>	SIVPYETNNQETFGALD	180
Qy	158	NYPDFOGARITITGOEGAGWITITNLL	LGKFSQKTRMF <td>SIVPYETNNQETFGALD</td> <td>217</td>	SIVPYETNNQETFGALD	217
Db	181	STQVTFVPONQOTIESPDNALQFRL	XGKRYNYNTHSFLCYGKDOAL	MOXLANDIOVASNEI	240
Qy	218	STQVTFVPONQOTIESPDNALQFRL	XGKRYNYNTHSFLCYGKDOAL	MOXLANDIOVASNEI	277
Db	241	LADPCFHHGYYKKVYVWSDLTKPTCT	KREEMTLPEQOFEIGIGVNOOC	HOSEILLEFNTSY	300
Qy	278	LADPCFHHGYYKKVYVWSDLTKPTCT	KREEMTLPEQOFEIGIGVNOOC	HOSEILLEFNTSY	337
Db	301	CPYSQCAFENGJFLPPLQGDGFA	FSAFYVVKFNLITSEKVSQEKVTE	MMKRFCAQPMEEI	360
Qy	338	CPYSQCAFENGJFLPPLQGDGFA	FSAFYVVKFNLITSEKVSQEKVTE	MMKRFCAQPMEEI	397
Db	361	KTSVAGVKEKXLTSECFSGTYILSL	LLQGYHFTADSWEHIFIKIGSDG	GMTLGYMLN	420
Qy	398	KTSVAGVKEKXLTSECFSGTYILSL	LLQGYHFTADSWEHIFIKIGSDG	GMTLGYMLN	457
Db	421	LTNMIPAEOPPLSTPLSHST	439		
Qy	458	LTNMIPAEOPPLSTPLSHST	476		
RESULT	7				
ID	PCT-US99-22955-6	STANDARD;	PRT;	454	AA.
AC	xxxxxx				
XX					
DT					
XX					
DE					
XX					
Sequence 6, Application PC/TUS9922955					
CC	Sequence 6, Application PC/TUS9922955				
CC	GENERAL INFORMATION:				
CC	APPLICANT: Maliszewski, Charles R.				
CC	APPLICANT: Gayle III, Richard B.				
CC	APPLICANT: Price, Virginia L.				
CC	APPLICANT: Gimpel, Steven D.				
CC	APPLICANT: Immunex Corporation				
CC	TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment				
CC	FILE REFERENCE: 2879-WO				
CC	CURRENT APPLICATION NUMBER: PCT/US99/22955				
CC	CURRENT FILING DATE: 1999-10-13				
CC	EARLIER APPLICATION NUMBER: US 60/104,585				
CC	EARLIER FILING DATE: 1998-10-16				
CC	EARLIER APPLICATION NUMBER: US 60/107,466				
CC	EARLIER FILING DATE: 1998-11-06				
CC	EARLIER APPLICATION NUMBER: US 60/149,010				
CC	EARLIER FILING DATE: 1999-08-13				
CC	NUMBER OF SEQ ID NOS: 31.				
CC	SOFTWARE: Patent In Ver. 2.0				
CC	SEQ ID NO 6				
CC	LENGTH: 454				
CC	TYPE: PRT				
CC	ORGANISM: Artificial Sequence				
CC	SEQUENCE 454 AA: 51543 MW: 1133800 CN:				
Query Match	86.0%; Score 3275; DB 1; Length 454;				
Best Local Similarity 100.0%; Pred. No. 0.00e+00;					
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0.					
Db	16	TQNKALPENVKYGIYLDAGSSHTSLYIK	WPAEKENDTGYYHOVEECRVKPGISK	VQK 75	
Qy	38	TQNKALPENVKYGIYLDAGSSHTSLYIK	WPAEKENDTGYYHOVEECRVKPGISK	VQK 97	
Db	76	VNEIGITVITDCMERARREVI	PPRSQHOETPVYLGATATGMLLM	MESEBELADRYLDVVERSL	135
Qy	98	VNEIGITVITDCMERARREVI	PPRSQHOETPVYLGATATGMLLM	MESEBELADRYLDVVERSL	157

Db 136 NYPDFOGARITIGQEGAGWITINLLGKFSQKTRWFSTVPEETNNQETFGALDLGGA 195
|
Qy 158 NYPDFOGARITIGQEGAGWITINLLGKFSQKTRWFSTVPEETNNQETFGALDLGGA 217
|
Db 196 STQVTFPQONQTIESPNNALQFRLYKDYNYTHSFCLCYGKDALMOKLAKDIOVASNEI 255
|
Qy 218 STQVTFPQONQTIESPNNALQFRLYKDYNYTHSFCLCYGKDALMOKLAKDIOVASNEI 277
|
Db 256 LRDPCHFPGYKRVVNSDLYKTPCTKRFEMLTPQOFEIOLGIGNYQOCHOSILELFNTSY 315
|
Qy 278 LRDPCHFPGYKRVVNSDLYKTPCTKRFEMLTPQOFEIOLGIGNYQOCHOSILELFNTSY 337
|
Db 316 CPYSQCAFNGIFLPPLOGDGAFAFVFWKFLNLITSEKVSQEKVTEMMKFCQAPWEEI 375
|
Qy 338 CPYSQCAFNGIFLPPLOGDGAFAFVFWKFLNLITSEKVSQEKVTEMMKFCQAPWEEI 397
|
Db 376 KTSYAGVKEKYLSEYCSFGTYILSLLOGYHFTADSMEHIFIGKIOGSDAGWTLGYMLN 435
|
Qy 398 KTSYAGVKEKYLSEYCSFGTYILSLLOGYHFTADSMEHIFIGKIOGSDAGWTLGYMLN 457
|
Db 436 LTNMIPAEOPLSTPLSHST 454
|
Qy 458 LTNMIPAEOPLSTPLSHST 476

RESULT 8
ID PCT-US99-22955-30 STANDARD; PRT; 463 AA.

AC xxxxxx

Sequence 30, Application PC/TUS9922955

CC GENERAL INFORMATION:
CC APPLICANT: Maliszewski, Charles R.
CC APPLICANT: Gayle III, Richard B.
CC APPLICANT: Price, Virginia L.
CC APPLICANT: Gimpel, Steven D.
CC APPLICANT: Immunex Corporation
CC TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment
CC FILE REFERENCE: 2879-WO
CC CURRENT APPLICATION NUMBER: PCT/US99/22955
CC CURRENT FILING DATE: 1999-10-13
CC EARLIER APPLICATION NUMBER: US 60/104,585
CC EARLIER FILING DATE: 1998-10-16
CC EARLIER APPLICATION NUMBER: US 60/107,466
CC EARLIER FILING DATE: 1998-11-06
CC EARLIER APPLICATION NUMBER: US 60/149,010
CC EARLIER FILING DATE: 1999-08-13
CC NUMBER OF SEQ ID NOS: 31
CC SOFTWARE: Patentln Ver. 2.0
CC SEQ ID NO 30
CC LENGTH: 463
CC TYPE: PRT
CC ORGANISM: Artificial Sequence
CC FEATURE:
CC OTHER INFORMATION: Description of Artificial Sequence: Fusion
CC OTHER INFORMATION: construct of human CD39
CC SEQUENCE 463 AA; 52621 MW; 1178782 CN;

Query Match 86.0%; Score 3275; DB 1; Length 463;

Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 25 TONKALPENYKYGIVLDAGSSHTSLYIKWPAEKENDTGVVHGYEBCRVKPGISKRYOK 84
|
Qy 38 TONKALPENYKYGIVLDAGSSHTSLYIKWPAEKENDTGVVHGYEBCRVKPGISKRYOK 97
|
Db 85 VNEIGIYLTDCMERAREVYIPRSQHOETPPVYLGATAGMRLRMESEELADRYLDVVERSL 144
|

Qy 98 VNEIGIYLTDCMERAREVYIPRSQHOETPPVYLGATAGMRLRMESEELADRYLDVVERSL 157
|
Db 145 NYPDFOGARITIGQEGAGWITINLLGKFSQKTRWFSTVPEETNNQETFGALDLGGA 204
|
Qy 158 NYPDFOGARITIGQEGAGWITINLLGKFSQKTRWFSTVPEETNNQETFGALDLGGA 217
|
Db 205 STQVTFPQONQTIESPNNALQFRLYKDYNYTHSFCLCYGKDALMOKLAKDIOVASNEI 264
|
Qy 218 STQVTFPQONQTIESPNNALQFRLYKDYNYTHSFCLCYGKDALMOKLAKDIOVASNEI 277
|
Db 265 LRDPCHFPGYKRVVNSDLYKTPCTKRFEMLTPQOFEIOLGIGNYQOCHOSILELFNTSY 324
|
Qy 278 LRDPCHFPGYKRVVNSDLYKTPCTKRFEMLTPQOFEIOLGIGNYQOCHOSILELFNTSY 337
|
Db 325 CPYSQCAFNGIFLPPLOGDGAFAFVFWKFLNLITSEKVSQEKVTEMMKFCQAPWEEI 384
|
Qy 338 CPYSQCAFNGIFLPPLOGDGAFAFVFWKFLNLITSEKVSQEKVTEMMKFCQAPWEEI 397
|
Db 385 KTSYAGVKEKYLSEYCSFGTYILSLLOGYHFTADSMEHIFIGKIOGSDAGWTLGYMLN 444
|
Qy 398 KTSYAGVKEKYLSEYCSFGTYILSLLOGYHFTADSMEHIFIGKIOGSDAGWTLGYMLN 457
|
Db 445 LTNMIPAEOPLSTPLSHST 463
|
Qy 458 LTNMIPAEOPLSTPLSHST 476

RESULT 9
ID PCT-US99-22955-29 STANDARD; PRT; 473 AA.

AC xxxxxx

Sequence 29, Application PC/TUS9922955

CC GENERAL INFORMATION:
CC APPLICANT: Maliszewski, Charles R.
CC APPLICANT: Gayle III, Richard B.
CC APPLICANT: Price, Virginia L.
CC APPLICANT: Gimpel, Steven D.
CC APPLICANT: Immunex Corporation
CC TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment
CC FILE REFERENCE: 2879-WO
CC CURRENT APPLICATION NUMBER: PCT/US99/22955
CC CURRENT FILING DATE: 1999-10-13
CC EARLIER APPLICATION NUMBER: US 60/104,585
CC EARLIER FILING DATE: 1998-10-16
CC EARLIER APPLICATION NUMBER: US 60/107,466
CC EARLIER FILING DATE: 1998-11-06
CC EARLIER APPLICATION NUMBER: US 60/149,010
CC EARLIER FILING DATE: 1999-08-13
CC NUMBER OF SEQ ID NOS: 31
CC SOFTWARE: Patentln Ver. 2.0
CC SEQ ID NO 29
CC LENGTH: 473
CC TYPE: PRT
CC ORGANISM: Artificial Sequence
CC FEATURE:
CC OTHER INFORMATION: Description of Artificial Sequence: Fusion
CC OTHER INFORMATION: construct of human CD39
CC SEQUENCE 473 AA; 53745 MW; 1229492 CN;

Query Match 86.0%; Score 3275; DB 1; Length 473;

Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 35 TONKALPENYKYGIVLDAGSSHTSLYIKWPAEKENDTGVVHGYEBCRVKPGISKRYOK 94
|
Qy 38 TONKALPENYKYGIVLDAGSSHTSLYIKWPAEKENDTGVVHGYEBCRVKPGISKRYOK 97
|
Db 95 VNEIGIYLTDCMERAREVYIPRSQHOETPPVYLGATAGMRLRMESEELADRYLDVVERSL 154
|

QY 98 VNEIGIYLTDCMERAREVIRPSOHOETPVYLGATAGMRLRMESEBELADRVLDVVERSL 157
DB 155 NYPDFOGARIIITGOEAGYGMWITINYLGRFSQKTRMFSTIYPETNNQETFGALDLGA 214
QY 158 NYPDFOGARIIITGOEAGYGMWITINYLGRFSQKTRMFSTIYPETNNQETFGALDLGA 217
DB 215 STQVTFVPONOTIESPDNALQFRLGKDYNYVTHSFCLCYGKQALMOKLAKDIOVASNEI 274
QY 218 STQVTFVPONOTIESPDNALQFRLGKDYNYVTHSFCLCYGKQALMOKLAKDIOVASNEI 277
DB 275 LRDPCHFHGPKYKVVNSDLKTPCTKREEMTLPFQOFIEIOGIGNYQOCHQSILELFTNSY 334
QY 278 LRDPCHFHGPKYKVVNSDLKTPCTKREEMTLPFQOFIEIOGIGNYQOCHQSILELFTNSY 337
DB 335 CPYSOCANFGITLPLQGDGFAFSAFYVMKFLNLTSEKVSQEKYTEMMKKFCQAOPMEI 394
QY 338 CPYSOCANFGITLPLQGDGFAFSAFYVMKFLNLTSEKVSQEKYTEMMKKFCQAOPMEI 397
DB 395 KTSYAGVKEKYLSEYCFSGTYILSLLOGYHFTADSMWHIHFICKIQGSDAGMTLGYMLN 454
QY 398 KTSYAGVKEKYLSEYCFSGTYILSLLOGYHFTADSMWHIHFICKIQGSDAGMTLGYMLN 457
DB 455 LTNMIPAEQPLSTPLSHST 473
QY 458 LTNMIPAEQPLSTPLSHST 476

RESULT 10
ID PCT-US99-22955-28 STANDARD: PRT: 474 AA.
XX
AC xxxxxx
XX
DT
XX
DE Sequence 28, Application PC/TUS9922955
CC
CC GENERAL INFORMATION:
CC APPLICANT: Maliszewski, Charles R.
CC APPLICANT: Gayle III, Richard B.
CC APPLICANT: Price, Virginia L.
CC APPLICANT: Gimpel, Steven D.
CC APPLICANT: Immunex Corporation
CC TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment
CC FILE REFERENCE: 2879-WO
CC CURRENT APPLICATION NUMBER: PCT/US99/22955
CC CURRENT FILING DATE: 1999-10-13
CC EARLIER APPLICATION NUMBER: US 60/104,585
CC EARLIER FILING DATE: 1998-10-16
CC EARLIER APPLICATION NUMBER: US 60/107,466
CC EARLIER FILING DATE: 1998-11-06
CC EARLIER APPLICATION NUMBER: US 60/149,010
CC EARLIER FILING DATE: 1999-08-13
CC NUMBER OF SEQ ID NOS: 31
CC SOFTWARE: Patentln Ver. 2.0
CC SEQ ID NO 28
CC LENGTH: 474
CC TYPE: PRT
CC ORGANISM: Artificial Sequence
CC FEATURE:
CC OTHER INFORMATION: Description of Artificial Sequence: Fusion
CC OTHER INFORMATION: construct of human CD39
CC SEQUENCE 474 AA: 53816 MW: 1234400 CN:

Query Match 86.0%; Score 3275; DB 1; Length 474;

Best Local Similarity 100.0%; Pred. No. 0.00e+00;

Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 36 TONKALPENVKYGIYLDAGSSHTSLIYKMPAEKENDTGVAHVQVECKRGKSGISKFYOK 95
QY 38 TONKALPENVKYGIYLDAGSSHTSLIYKMPAEKENDTGVAHVQVECKRGKSGISKFYOK 97

DB 96 VNEIGIYLTDCMERAREVIRPSOHOETPVYLGATAGMRLRMESEBELADRVLDVVERSL 155
QY 98 VNEIGIYLTDCMERAREVIRPSOHOETPVYLGATAGMRLRMESEBELADRVLDVVERSL 157
DB 156 NYPDFOGARIIITGOEAGYGMWITINYLGRFSQKTRMFSTIYPETNNQETFGALDLGA 215
QY 158 NYPDFOGARIIITGOEAGYGMWITINYLGRFSQKTRMFSTIYPETNNQETFGALDLGA 217
DB 216 STQVTFVPONOTIESPDNALQFRLGKDYNYVTHSFCLCYGKQALMOKLAKDIOVASNEI 275
QY 218 STQVTFVPONOTIESPDNALQFRLGKDYNYVTHSFCLCYGKQALMOKLAKDIOVASNEI 277
DB 276 LRDPCHFHGPKYKVVNSDLKTPCTKREEMTLPFQOFIEIOGIGNYQOCHQSILELFTNSY 335
QY 278 LRDPCHFHGPKYKVVNSDLKTPCTKREEMTLPFQOFIEIOGIGNYQOCHQSILELFTNSY 337
DB 336 CPYSOCANFGITLPLQGDGFAFSAFYVMKFLNLTSEKVSQEKYTEMMKKFCQAOPMEI 395
QY 338 CPYSOCANFGITLPLQGDGFAFSAFYVMKFLNLTSEKVSQEKYTEMMKKFCQAOPMEI 397
DB 396 KTSYAGVKEKYLSEYCFSGTYILSLLOGYHFTADSMWHIHFICKIQGSDAGMTLGYMLN 455
QY 398 KTSYAGVKEKYLSEYCFSGTYILSLLOGYHFTADSMWHIHFICKIQGSDAGMTLGYMLN 457
DB 456 LTNMIPAEQPLSTPLSHST 474
QY 458 LTNMIPAEQPLSTPLSHST 476

RESULT 11
ID PCT-US99-22955-3 STANDARD: PRT: 476 AA.
XX
AC xxxxxx
XX
DT
XX
DE Sequence 3, Application PC/TUS9922955
CC
CC GENERAL INFORMATION:
CC APPLICANT: Maliszewski, Charles R.
CC APPLICANT: Gayle III, Richard B.
CC APPLICANT: Price, Virginia L.
CC APPLICANT: Gimpel, Steven D.
CC APPLICANT: Immunex Corporation
CC TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment
CC FILE REFERENCE: 2879-WO
CC CURRENT APPLICATION NUMBER: PCT/US99/22955
CC CURRENT FILING DATE: 1999-10-13
CC EARLIER APPLICATION NUMBER: US 60/104,585
CC EARLIER FILING DATE: 1998-10-16
CC EARLIER APPLICATION NUMBER: US 60/107,466
CC EARLIER FILING DATE: 1998-11-06
CC EARLIER APPLICATION NUMBER: US 60/149,010
CC EARLIER FILING DATE: 1999-08-13
CC NUMBER OF SEQ ID NOS: 31
CC SOFTWARE: Patentln Ver. 2.0
CC SEQ ID NO 3
CC LENGTH: 476
CC TYPE: PRT
CC ORGANISM: Artificial Sequence
CC FEATURE:
CC OTHER INFORMATION: Description of Artificial Sequence: Fusion
CC OTHER INFORMATION: construct of human CD39
CC SEQUENCE 476 AA: 54177 MW: 1244820 CN:

Query Match 86.0%; Score 3275; DB 1; Length 476;

Best Local Similarity 100.0%; Pred. No. 0.00e+00;

Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 38 TONKALPENVKYGIYLDAGSSHTSLIYKMPAEKENDTGVAHVQVECKRGKSGISKFYOK 97
QY 38 TONKALPENVKYGIYLDAGSSHTSLIYKMPAEKENDTGVAHVQVECKRGKSGISKFYOK 97

Db 98 VNEIGIYLTDCMERAREVIRRSQHETPVYLGATAGMRLLRMESEELADRYLDVVERSL 157
|
Qy 98 VNEIGIYLTDCMERAREVIRRSQHETPVYLGATAGMRLLRMESEELADRYLDVVERSL 157
|
Db 158 NYPEDFOGARITIGOEAGAWTINTYLLGKFSQKTRMFSTIVPEYETNNQETFGALDLGA 217
|
Qy 158 NYPEDFOGARITIGOEAGAWTINTYLLGKFSQKTRMFSTIVPEYETNNQETFGALDLGA 217
|
Db 218 STQYTFVPQONTIESPNDALQFRLYGKDYNYTHSFCLCYGKDQALMOKLAKDIOVASNEI 277
|
Qy 218 STQYTFVPQONTIESPNDALQFRLYGKDYNYTHSFCLCYGKDQALMOKLAKDIOVASNEI 277
|
Db 278 LRDCPFHFGKRYVNSDLKTKPCTKRFEMLTPQOEIIGIGNYOCHOSILELFNTSY 337
|
Qy 278 LRDCPFHFGKRYVNSDLKTKPCTKRFEMLTPQOEIIGIGNYOCHOSILELFNTSY 337
|
Db 338 CPYSQCAFNGIFLPLDGDGAFSAFYVVKFNLJTSEKVSQEKVTBMKKFCAQPMEEI 397
|
Qy 338 CPYSQCAFNGIFLPLDGDGAFSAFYVVKFNLJTSEKVSQEKVTBMKKFCAQPMEEI 397
|
Db 398 KTSYAGVKEKYLSEYCSGYIYLSLLQGYHFTADSWEHIFIGKIOGSDAGWTLYMYLN 457
|
Qy 398 KTSYAGVKEKYLSEYCSGYIYLSLLQGYHFTADSWEHIFIGKIOGSDAGWTLYMYLN 457
|
Db 458 LTNNMIPAEOPLSTPLSHST 476
|
Qy 458 LTNNMIPAEOPLSTPLSHST 476

RESULT 12
ID PCT-US99-22955-8 STANDARD: PRT: 478 AA.

AC xxxxxx

Sequence 8, Application PC/TUS9922955

CC Sequence 8, Application PC/TUS9922955
CC GENERAL INFORMATION:
CC APPLICANT: Maliszewski, Charles R.
CC APPLICANT: Gayle III, Richard B.
CC APPLICANT: Price, Virginia L.
CC APPLICANT: Gimpel, Steven D.
CC APPLICANT: Immunex Corporation
CC TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment
CC FILE REFERENCE: 2879-WO
CC CURRENT APPLICATION NUMBER: PCT/US99/22955
CC EARLIER FILING DATE: 1999-10-13
CC EARLIER APPLICATION NUMBER: US 60/104,585
CC EARLIER FILING DATE: 1998-10-16
CC EARLIER APPLICATION NUMBER: US 60/107,466
CC EARLIER FILING DATE: 1998-11-06
CC EARLIER APPLICATION NUMBER: US 60/149,010
CC EARLIER FILING DATE: 1999-08-13
CC NUMBER OF SEQ ID NOS: 31
CC SOFTWARE: Patentln Ver. 2.0
CC SEQ ID NO 8
CC LENGTH: 478
CC TYPE: PRT
CC ORGANISM: Artificial Sequence
CC SEQUENCE 478 AA; 54188 MW; 1255693 CN;

Query Match 86.0%; Score 3275; DB 1; Length 478;

Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 40 TONKALPENVKYGIYLDAGSSHTSLYTKWPAEKENDTGVAHVQECRVKPGISKFYOK 99
|
Qy 38 TONKALPENVKYGIYLDAGSSHTSLYTKWPAEKENDTGVAHVQECRVKPGISKFYOK 97
|
Db 100 VNEIGIYLTDCMERAREVIRRSQHETPVYLGATAGMRLLRMESEELADRYLDVVERSL 157

Qy 98 VNEIGIYLTDCMERAREVIRRSQHETPVYLGATAGMRLLRMESEELADRYLDVVERSL 157
|
Db 160 NYPEDFOGARITIGOEAGAWTINTYLLGKFSQKTRMFSTIVPEYETNNQETFGALDLGA 219
|
Qy 158 NYPEDFOGARITIGOEAGAWTINTYLLGKFSQKTRMFSTIVPEYETNNQETFGALDLGA 217
|
Db 220 STQYTFVPQONTIESPNDALQFRLYGKDYNYTHSFCLCYGKDQALMOKLAKDIOVASNEI 279
|
Qy 218 STQYTFVPQONTIESPNDALQFRLYGKDYNYTHSFCLCYGKDQALMOKLAKDIOVASNEI 277
|
Db 280 LRDCPFHFGKRYVNSDLKTKPCTKRFEMLTPQOEIIGIGNYOCHOSILELFNTSY 339
|
Qy 278 LRDCPFHFGKRYVNSDLKTKPCTKRFEMLTPQOEIIGIGNYOCHOSILELFNTSY 337
|
Db 340 CPYSQCAFNGIFLPLDGDGAFSAFYVVKFNLJTSEKVSQEKVTBMKKFCAQPMEEI 399
|
Qy 338 CPYSQCAFNGIFLPLDGDGAFSAFYVVKFNLJTSEKVSQEKVTBMKKFCAQPMEEI 397
|
Db 400 KTSYAGVKEKYLSEYCSGYIYLSLLQGYHFTADSWEHIFIGKIOGSDAGWTLYMYLN 459
|
Qy 398 KTSYAGVKEKYLSEYCSGYIYLSLLQGYHFTADSWEHIFIGKIOGSDAGWTLYMYLN 457
|
Db 460 LTNNMIPAEOPLSTPLSHST 478
|
Qy 458 LTNNMIPAEOPLSTPLSHST 476

RESULT 13
ID PCT-US99-22955-26 STANDARD: PRT: 487 AA.

AC xxxxxx

Sequence 26, Application PC/TUS9922955

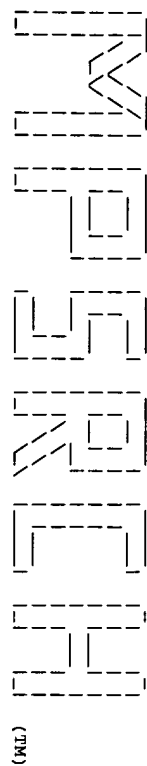
CC Sequence 26, Application PC/TUS9922955
CC GENERAL INFORMATION:
CC APPLICANT: Maliszewski, Charles R.
CC APPLICANT: Gayle III, Richard B.
CC APPLICANT: Price, Virginia L.
CC APPLICANT: Gimpel, Steven D.
CC APPLICANT: Immunex Corporation
CC TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment
CC FILE REFERENCE: 2879-WO
CC CURRENT APPLICATION NUMBER: PCT/US99/22955
CC EARLIER FILING DATE: 1999-10-13
CC EARLIER APPLICATION NUMBER: US 60/104,585
CC EARLIER FILING DATE: 1998-10-16
CC EARLIER APPLICATION NUMBER: US 60/107,466
CC EARLIER FILING DATE: 1998-11-06
CC EARLIER APPLICATION NUMBER: US 60/149,010
CC EARLIER FILING DATE: 1999-08-13
CC NUMBER OF SEQ ID NOS: 31
CC SOFTWARE: Patentln Ver. 2.0
CC SEQ ID NO 26
CC LENGTH: 487
CC TYPE: PRT
CC ORGANISM: Artificial Sequence
CC SEQUENCE 487 AA; 55240 MW; 1301482 CN;

Query Match 86.0%; Score 3275; DB 1; Length 487;

Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 49 TONKALPENVKYGIYLDAGSSHTSLYTKWPAEKENDTGVAHVQECRVKPGISKFYOK 108
|
Qy 38 TONKALPENVKYGIYLDAGSSHTSLYTKWPAEKENDTGVAHVQECRVKPGISKFYOK 97
|
Db 109 VNEIGIYLTDCMERAREVIRRSQHETPVYLGATAGMRLLRMESEELADRYLDVVERSL 168
|
Qy 98 VNEIGIYLTDCMERAREVIRRSQHETPVYLGATAGMRLLRMESEELADRYLDVVERSL 157

0Y	49	YGIYLDGSSHTSLYIKKMPAKENDTGVMHQVEBCRVKPGISKFVQKVNIGIYITDC	108
Db	109	MERAREVTPRSSQOHEPVPYLGAIAAMRLRLRNESELADRVLVDVENSLSNYPFDQGARI	168
QY	109	MERAREVTPRSSQOHEPVPYLGAIAAMRLRLRNESELADRVLVDVENSLSNYPFDQGARI	168
Db	169	ITGGEBAVWMTINTLILKGFQSKTRMFSLYPTETNNQETFGALDGGASIVTVPQNO	228
QY	169	ITGGEBAVWMTINTLILKGFQSKTRMFSLYPTETNNQETFGALDGGASIVTVPQNO	228
Db	229	TIESPDNALOFRLYGKDYNVYTHSFLCYGKDQALMOKLANDIQVANSIELLRDPCFHPGYK	288
QY	229	TIESPDNALOFRLYGKDYNVYTHSFLCYGKDQALMOKLANDIQVANSIELLRDPCFHPGYK	288
Db	289	KVWVNSLYTPTCKTFEWTLPFOOFELQIGIGNQOCHOSILELFTPTSCIPSCQAFNGI	348
QY	289	KVWVNSLYTPTCKTFEWTLPFOOFELQIGIGNQOCHOSILELFTPTSCIPSCQAFNGI	348
Db	349	FLPPLQODFCAFSAFVFKFNLNSEVSOEKTENMKKFCQAQPEELKTSYAGVREKY	408
QY	349	FLPPLQODFCAFSAFVFKFNLNSEVSOEKTENMKKFCQAQPEELKTSYAGVREKY	408
Db	409	LSEYCESGTYIISLLQGYHFTADSWEHIFRIGKIQGSDAGWTLGMYMLTNTMTIPAEQPL	468
QY	409	LSEYCESGTYIISLLQGYHFTADSWEHIFRIGKIQGSDAGWTLGMYMLTNTMTIPAEQPL	468
Db	469	STPLSHT 476	
QY	469	STPLSHT 476	
RESULT	15		
ID	US-08-701-460-4	STANDARD;	PRT; 510 AA.
XX	xxxxxx		
DE	Sequence 4, Application US/08701460		
CC	GENERAL INFORMATION:		
CC	APPLICANT: MARCUS, AARON J.		
CC	APPLICANT: MALISZEWSKI, CHARLES R.		
CC	TITLE OF INVENTION: METHODS OF REGULATING HEMOSTASIS AND		
CC	TITLE OF INVENTION: THROMBOSIS, AND COMPOUNDS USEFUL THEREFOR		
CC	NUMBER OF SEQUENCES: 16		
CC	CORRESPONDENCE ADDRESSES:		
CC	ADDRESSEE: Immunex Corporation		
CC	STREET: 51 University Street		
CC	CITY: Seattle		
CC	STATE: WA		
CC	COUNTRY: USA		
CC	ZIP: 98101		
CC	COMPUTER READABLE FORM:		
CC	MEDIUM TYPE: Floppy disk		
CC	COMPUTER: Apple Power Macintosh 7200/90		
CC	OPERATING SYSTEM: Apple Operating System 7.5.3		
CC	SOFTWARE: Microsoft Word 6.0.1 for Power Macintosh		
CC	CURRENT APPLICATION DATA:		
CC	APPLICATION NUMBER: US/08/701,460		
CC	FILING DATE: 22 AUGUST 1996		
CC	CLASSIFICATION: 514		
CC	ATTORNEY/AGENT INFORMATION:		
CC	NAME: Perkins, Patricia Anne		
CC	REGISTRATION NUMBER: 34,693		
CC	REFERENCE/DOCKET NUMBER: 2808		
CC	TELECOMMUNICATION INFORMATION:		
CC	TELEPHONE: 206-587-0430		
CC	INFORMATION FOR SEQ ID NO: 4:		
CC	SEQUENCE CHARACTERISTICS:		
CC	LENGTH: 510 amino acids		



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution Rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Fri May 5 08:41:38 2000; Maspar time 49.00 Seconds
Tabular output not generated. 490.937 Million cell updates/sec

Title: >US-09-374-586-1
Description: (1-510) from US09374586.pep
Perfect Score: 3806
Sequence: 1 MEDTKESNVKTFCSKNILAI.....VAITGLLPHKPSYWKDNY 510

Scoring table:
PAM 150
Gap 11

Searched: 142080 seqs, 47172406 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pif62
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 51.082; Variance 106.884; scale 0.478
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	3806	100.0	510	2	I56242 lymphoid cell activat	0.00e+00
2	378	9.9	537	2	T16696 hypothetical protein	4.12e-44
3	360	9.5	630	2	S50463 hypothetical protein	7.21e-41
4	319	8.4	454	2	JC4616 apyrase (EC 3.6.1.5)	1.40e-33
5	293	7.7	1052	2	T04439 hypothetical protein	4.85e-29
6	291	7.6	455	2	S48859 nucleoside triphospha	1.08e-28
7	265	7.0	518	2	A40732 guanosine diphospha	3.06e-24
8	131	3.4	371	2	T05213 hypothetical protein	1.75e-03
9	125	3.3	597	2	T04954 hypothetical protein	1.06e-02
10	127	3.1	628	2	A55421 nucleoside-triphosha	5.84e-03
11	117	3.1	528	2	JN0620 UDP-glucuronosyltrans	1.07e-01
12	116	3.0	107	2	G69988 thioredoxin H1 homo	1.42e-01
13	113	3.0	919	2	S37786 hypothetical protein	3.29e-01
14	114	3.0	974	1	EYLPND H+-transporting ATPas	2.49e-01
15	110	2.9	369	2	A72453 conserved hypothetical	5.70e-01
16	111	2.9	392	2	A72453 conserved hypothetical	7.48e-01
17	109	2.9	586	2	T12886 conserved hypothetical	9.81e-01
18	106	2.8	204	2	C58930 NADH dehydrogenase (u	2.19e+00
19	107	2.8	276	2	S13585 most protein precurs	1.68e+00
20	108	2.8	466	2	PC4296 nicotinic acetylcholi	1.28e+00
21	106	2.8	521	2	C28529 nicotinic acetylcholi	2.19e+00
22	106	2.8	573	2	F70860 probable cytochrome-c	2.19e+00
23	108	2.8	669	2	T05212 hypothetical protein	1.28e+00

24	104	2.7	44	2	S63501 apyrase (EC 3.6.1.5)	3.69e+00
25	101	2.7	222	2	C70020 conserved hypothetical	7.99e+00
26	104	2.7	296	2	A72239 conserved hypothetical	3.69e+00
27	102	2.7	450	2	JN0105 tubulin alpha-3 chain	6.19e+00
28	103	2.7	493	1	ACRYB1 nicotinic acetylcholi	4.79e+00
29	102	2.7	494	2	T03774 probable histidine--t	6.19e+00
30	101	2.7	501	2	S04607 nicotinic acetylcholi	7.99e+00
31	101	2.7	505	2	S07227 nicotinic acetylcholi	7.99e+00
32	102	2.7	530	2	A48633 glucuronosyltransfera	6.19e+00
33	104	2.7	544	2	T05211 hypothetical protein	3.69e+00
34	101	2.7	557	2	S12359 nicotinic acetylcholi	7.99e+00
35	103	2.7	571	1	RNCW7H transcription initiat	4.79e+00
36	101	2.7	571	1	RNCW7H transcription initiat	7.99e+00
37	102	2.7	574	2	T05214 hypothetical protein	6.19e+00
38	102	2.7	634	2	S75962 NADH dehydrogenase (u	6.19e+00
39	103	2.7	787	2	S72725 guanosine-3',5'-bis(d	4.79e+00
40	102	2.7	790	2	F70725 probable rela protein	6.19e+00
41	103	2.7	867	2	G69485 DNA-directed RNA poly	4.79e+00
42	102	2.7	923	2	S44226 periodic tryptophan p	6.19e+00
43	104	2.7	1680	2	T01367 hypothetical protein	7.99e+00
44	101	2.7	2672	2	A48126 translation activator	3.69e+00
45	100	2.6	461	2	S68137 NADH dehydrogenase (u	1.03e+01

ALIGNMENTS

RESULT	1	I56242	#type complete	lymphoid cell activation antigen - human
ENTRY				#formal_name Homo sapiens #common_name man
TITLE				02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change
ORGANISM				02-Jul-1996
DATE				I56242
ACCESSIONS				I56242
REFERENCE				Maliszewski, C.R.; Delespess, G.J.; Schoenborn, M.A.;
#authors				Armitage, R.J.; Fanslow, W.C.; Nakajima, T.; Baker, E.;
				Sutherland, G.R.; Polindexter, K.; Birks, C.
#journal				J. Immunol. (1994) 153:3574-3583
#title				The CD39 lymphoid cell activation antigen. Molecular cloning
				and structural characterization.
#cross-references				MUID:95015846
#accession				I56242
#status				preliminary; translated from GB/EMBL/DBJ
#molecule_type				mRNA
#residues				1-510 #label RES
#cross-references				GB:S7813; NID:9765255; PID:9765256
SUMMARY				#length 510 #molecular_weight 57964 #checksum 4387
Query Match				100.0%; Score 3806; DB 2; Length 510;
Best Local Similarity				100.0%; Pred. No. 0.00e+00;
Matches				510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db	1	MEDTKESNVKTFCSKNILAI	IGFSSIIAVIALAVGLTQNKALPENKYGIVLDAGSSHT	60
Qy	1	MEDTKESNVKTFCSKNILAI	IGFSSIIAVIALAVGLTQNKALPENKYGIVLDAGSSHT	60
Db	61	SLTYIKWPAEKENDTGVHVECECRVKGPSIKSFVQVKNIGIYLTDCMERAREVIPSQ	120	
Qy	61	SLTYIKWPAEKENDTGVHVECECRVKGPSIKSFVQVKNIGIYLTDCMERAREVIPSQ	120	
Db	121	HOETPVYLGATAGRLLRMESEELADRVLDVVERSSINYPDEFGARITIGQEGAGWI	180	
Qy	121	HOETPVYLGATAGRLLRMESEELADRVLDVVERSSINYPDEFGARITIGQEGAGWI	180	
Db	181	TINTLAKFSQKTRMFSIVPEYETNNOETFGALDGGASTQVTFPQNOTIESPNALOFR	240	
Qy	181	TINTLAKFSQKTRMFSIVPEYETNNOETFGALDGGASTQVTFPQNOTIESPNALOFR	240	
Db	241	LYGKDYVYTHSFTCYGKDALMOKLAKDIOVASNEILRDPCEHPGKXKVVNSVLDYKTP	300	
Qy	241	LYGKDYVYTHSFTCYGKDALMOKLAKDIOVASNEILRDPCEHPGKXKVVNSVLDYKTP	300	
Db	301	CTKFEKTLPRQPEFEGISGNYOCHOSILELFTWISICPSQCAFNGILFLPRLQGDGCAP	360	

QY	301	CTKREMTLPQGFPIQIGINVCQCHOSILBLEFNSTYPSYQCAFNGTFLPPLQDGCAG	360
Db	361	SAFYFVMKFLNLTSEKVSQEKVTEMMKRFCAQPMEEIKTSTAGVKEKYLSEYCSGTIYL	420
QY	361	SAFYFVMKFLNLTSEKVSQEKVTEMMKRFCAQPMEEIKTSTAGVKEKYLSEYCSGTIYL	420
Db	421	SLLLQGVHPTADSWHIFHFIKIGQSDAGWTLGWLNTLNMIPAEQPLSTPLSHSTYVFL	480
QY	421	SLLLQGVHPTADSWHIFHFIKIGQSDAGWTLGWLNTLNMIPAEQPLSTPLSHSTYVFL	480
Db	481	MYLFLSVLFTVAIIGLLIFHFKPSYFMRKMV	510
QY	481	MYLFLSVLFTVAIIGLLIFHFKPSYFMRKMV	510
RESULT	2		
ENTRY	T16696	#type complete	
TITLE	hypothetical protein R07E4.4 - <i>Caenorhabditis elegans</i>		
ORGANISM	#formal_name <i>Caenorhabditis elegans</i>		
DATE	20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999		
ACCESSIONS	T16696		
REFERENCE	T16696		
#authors	Miller, N.		
#submission	submitted to the EMBL Data Library, October 1995		
#description	The sequence of <i>C. elegans</i> cosmid R07E4.		
#accession	T16696		
#status	preliminary; translated from GB/EMBL/DBJ		
#molecule_type	DNA		
#residues	1-557 #label MTL		
#cross-references	EMBL:U39652; NID:q1049390; PID:q1049394; PIDN:AAA80403.1; CESP:R07E4.4		
GENETICS	#experimental_source strain Bristol N2		
#gene	CESP:R07E4.4		
#introns	39/1; 67/1; 145/2; 192/1; 244/3; 317/3; 451/3; 488/2		
SUMMARY	#length 557 #molecular_weight 63056 #checksum 4267		
Query Match	9.9%; Score 378; DB 2; Length 557;		
Best Local Similarity	23.9%; Pred. NO. 4.12e-44;		
Matches	122; Conservative 134; Mismatches 217; Indels 37; Gaps 30;		
Db	12	VSAMIEFPVI-VETIYVEAHTSPKVIADDERSYGVICDAGSTGRFLFYNMISTDSEL	70
QY	18	LAILGFSSIIIVIALAVGLTQNKALPBN-VK-YGIYVADAGSSHTSLYIKWPAKEKNDI	75
Db	71	IQIEYIYDKNRPYMKKISIPGSLSTGCTKRAQAAEYLRPLMEALAEKHIEBEKRPYPTIFA	130
QY	76	GVIHQVE-ECR-V-K-GPGISKFPVQKNEIGITLDCMEARAREVIRPSOHOETPVYIGA	130
Db	131	TAGMRLDEVELYIGOKEAVALKLNLRNKLPTISMQVLKEHRIIEKGWEGIYSAIVNYA	190
QY	131	TAGMRLMEBEELADR-VLDVYERLSLN-PRD-FQG-ARITIGDEGAYIGMTITNYL	165
Db	191	LGFENKATATLDPGTSAPAHAKQTVGMIDMGASAAQIAFELPDTDSFSSINVENINLGR	250
QY	186	LGRFSQKTRM-FSIYVPEYTNNOETFGALDICASTQVTF-VPAQQTIES-P-DV-ALQFR	240
Db	251	EDDSLFKKKLVTFVFLGIVNEGIRKYEHLMSLKAKQNGVIVIDDCLPMLNKHVTYLEN	310
QY	241	LYGR-DVNYVTHSFLYCGKQOALM--QKIAKDIQVANSINLIDPCFHPQKRVVAVSD	295
Db	311	G-ENR-VRGTGNMNTCSNEVKLLN-PSSSEYVCKA-EAAKCYFAGVAPASIPLSNIEM	366
QY	296	LYKRPCTKRFEMTLRPFQGFPIQIGINVCQCHOSILBLEFNSTYPSYQCAFNGTFLPPLQD	355
Db	367	-YG-FSEWYSTHVDLYGGO-YDAENIAKKTQOYCSKRMSTIOAESKKOLYPRADERL	423
QY	356	DFGAFSAFYFVMK-FLNLTSEKVSQEKVTEMMKRFCAQPM--E-EIKTS-YAGVKEKYL	409
Db	424	RTQGCASMTISVLHDS-VDKTHNKFQSVSTTLAGOEYOMALGAMTYHNRFPPLDSSR	462

Yy	410	SEPCSSGYIILSLAQGHVFPADSMENHFGIKIQGSDAGWITGLYMLNLTMMIPAEQPLS	469
Db	483	NLIYKETHSSSESLMAPFLFSLAVFCFLV	512
Qy	470	TPLSHSTYVFLMVLFSLVLTFTVAIGLLIF	499
RESULT	3		
ENTRY		S50463	#type complete
TITLE		hypothetical protein YER005w - yeast (<i>Saccharomyces cerevisiae</i>)	
ORGANISM		#formal_name <i>Saccharomyces cerevisiae</i>	
DATE		28-May-1993	#sequence_revision 24-Feb-1995
		21-Nov-1997	#text_change
ACCESSIONS		S50463	
REFERENCE		S50428	
#authors		Dietrich, F.S.	
#submission		Submitted to the EMBL Data Library, December 1994	
#description		The sequence of <i>S. cerevisiae</i> cosmid 9537, 9581, 9495, 9867 and lambda clone 5898.	
#accession		S50463	
#molecule_type		DNA	
#residues		1-630	#label DIE
#cross_references		EMBL:018778; NID:g603592; PID:g603597; MIPS:YER005w	
GENETICS			
SUMMARY		#map_position 5R	
		length 630	#molecular_weight 71851
			#checksum 7317
Query Match		9.5%; Score 360; DB 2; Length 630;	
Best Local Similarity		28.4%; Pred. No. 7,21e-41;	
Matches		106; Conservative 93; Mismatches 137; Indels 37; Gaps 28;	
Db	52	THQEKDWTFKLNPGISLSEFKRPQDAYKSHIKPLIDFAKNIIPESHWSSCFVYIQATAGMR	111
Qy	78	VHOVEECRVK-GPGISKRVQKVNKEI-GIYLLDCKERAREVYIPRSQHOETPVYLGATAGMR	135
Db	112	LLPQDIQSLIDGLCGQGLKHP-AEFLVEDCSAQIQVYIDGTEBGLGMLGNYLYGHND-	169
Qy	136	LLRHESEE-LADRYLADVVERSLWYPR-DFQGA-RITTGEEGAYGITTNYLLGKRSQ	192
Db	170	--YN--P-EVSDHFTGFCFMDMGASTQIAPAPHDSEIARHRDITIFLRVSGDLQK	223
Qy	193	TRWFSIYFETNNQETGALDLGASTQVYTFVPONR-IESP-DN-ALQF-R-LYKQ-D-	245
Db	224	WDVSVTWLIGAGQAQRRRYLAQILNLPENT-ND-AENDDFSTR-NLNDPCAPRGSSSTD	280
Qy	246	YNYVTHSLKCGKQALMOKLAKIQVANSNLTMDPCFHRGKYKVVAVSDLYKTP-CTKR	304
Db	281	FEFKDTI--PHIAGSGNEQCTKSIYPLLLKNMPCDDPCLFNGVAPRIIDFANDKFIGT	338
Qy	305	FEMTLPEQOFELQIGIGNVOCHOSILELFTNYSY-CPYSOCAFNIGIEFLPQD-C-D-FGAF	360
Db	339	SEWYMTADVVKLGE-YNEDKFSKSLREPCNSWWTQILANSQKGVNSIPENFLDQAC	397
Qy	361	SAFFVPMK-FNLNLSEKVSQEKVTEMMKKFCQAPWEEI-----KTSTAGVKEYLSBICF	414
Db	398	KGNVNLNLHGEF	410
Qy	415	SGTYILSLILQGY	427
RESULT	4		
ENTRY		JC4616	#type complete
TITLE		apsyrase (EC 3.6.1.5) precursor - potato	
ALTERNATE_NAMES		adenylpyrophosphatase; ATP-diphosphohydrolase	
ORGANISM		#formal_name <i>Solanum tuberosum</i>	#common_name potato
DATE		10-May-1996	#sequence_revision 19-Jul-1996
		28-May-1999	#text_change
ACCESSIONS		JC4616; PC4147	
REFERENCE		JC4616	
#authors		Handa, M.; Guidotti, G.	
#journal		Biochem. Biophys. Res. Commun. (1996) 218:916-923	
#title		Purification and cloning of a soluble ATP-diphosphohydrolase	

(APYrase) from potato tubers (Solanum tuberosum).
#cross-references MUID:96158985
#accession J04616
##molecule-type mRNA
##residues 1-454 ##label HAN
##cross-references GB:058597; NID:91381632; PIDN:AMB02720.1;
PID:91381633

#accession PC4147
##molecule-type protein
##residues 59-95;96-131;132-160;236-253;332-345 ##label HA2
##experimental_source tubers
##note
Gly
The authors translated the codon GCA for residue 215 as Gly

COMMENT This enzyme belongs to a family of E-type APases, and it catalyzes the hydrolysis of phosphoanhydride bonds of nucleoside tri- and di- phosphates in the presence of divalent cations. This enzyme has nucleotide substrate specificity, divalent cation requirement, and insensitivity to known inhibitors of APases, phosphatases, and adenylate kinase. It is a regulator of various steps involved in starch synthesis.

GENETICS

CLASSIFICATION rtop1
#superfamily nucleoside triphosphatase chromatin-associated glycoprotein; hydrolase; transmembrane protein

KEYWORDS

FEATURE 1-30
8-25 #domain signal sequence #status predicted #label SIG
31-454 #domain transmembrane #status predicted #label TM
#product ATP-diphosphohydrolase #status predicted #label MAP

44-65 #region actin-heat shock protein 70-hexokinase
beta-phosphate binding
44-65 #region nucleotide binding #status predicted
192-212 #region actin-heat shock protein 70-hexokinase
gamma-phosphate binding
192-212 #region nucleotide binding #status predicted
390-410,427-446 #region hydrophobic carboxyl end
151,262 #binding site carbohydrate (asn) (covalent) #status predicted

SUMMARY #length 454 #molecular-weight 50040 #checksum 6727

Query Match 8.4%; Score 319; DB 2; Length 454;
Best Local Similarity 28.3%; Pred. No. 1,40e-33;
Matches 69; Conservative 66; Mismatches 99; Indels 10; Gaps 9;

Db 12 ILAFLVPLSLSKNNAQIPLRHLHSEHYAVIFDAGSGRYHFRF-DEKIGL 70
||||| :
17 ILAI-LGSS-SIIAVIALLAGLITONKALPENVKYGIILDGSSHTSLYIKWPAEKEND 74
||| :
71 LPIGNNTIEFMATPERGSSVAEDPKAANSLEPLLDGEGVVPQELQSEMPLEIGATGL 130
||| :
75 TGVVHOVEECVKPGSKFQKYNKIGIYITDCMERAREVYPSQHOETPVYLGATAGM 134
||| :
131 RMLGDAEAKILIAVRNLY-KNOSTFHSKDQWVITLDGTORGSSYMAAINTLGNLG-KD 188
| :
135 RLKMESE-LADVLDIVERSLSNYPDFGARIITGQEGAGAWITITINLLKFSQKT 193
| :
Db 189 --YKST-ATIDLG-GSVQMAVAISNEQAKAQONDEGEPEYQKHLMSKDYLVYHSY 244
| :
194 RWFSEIVETNNQETFGALDGGASTQVTEVPQQTIESPDNALQFRILYGRDYNVYHSF 253
| :
Db 245 LNYG 248
| :
QY 254 LNYG 257

RESULT 5
ENTRY T04439 #type complete
TITLE hypothetical protein T18B16.150 Arabidopsis thaliana
ORGANISM Arabidopsis thaliana #common_name mouse-ear
#formal_name Arabidopsis thaliana
#accessions T04439
#sequence_revision 30-Apr-1999 #text_change 30-Apr-1999

REFERENCE 215359
#authors Bevan, M.; Benes, V.; Rechmann, S.; Borkova, D.; Ansoorge, W.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Scheller, C.
#submission submitted to the Protein Sequence Database, April 1998
#accession T04439
##molecule-type DNA
##residues 1-1052 ##label BEV
##cross-references EMBL:AL021687
##experimental_source cultivar Columbia; BAC clone T18B16

GENETICS
#map_position 4
#introns 79/3; 118/3; 217/1; 295/3; 396/3; 531/2; 815/3
#note T18B16.150
SUMMARY #length 1052 #molecular-weight 116388 #checksum 4955

Query Match 7.7%; Score 293; DB 2; Length 1052;
Best Local Similarity 27.3%; Pred. No. 4.85e-29;
Matches 115; Conservative 120; Mismatches 135; Indels 52; Gaps 38;

Db 578 RGFPIVNNRTGLTKAIKPLQMAEKQIPKNAHRTSLFYVATAGVRLR--PAD-SSMI 634
||| :
89 FGISKFOKNEIGIYITDCMERAREVYPSQHOETPVYLGATAGMRLRMESELDARV 148
| :
Db 635 LGNWSILAKSPFCRRREVMKIIIGTEAYFGWTLANT-----QTSW-LGALP---KK- 683
| :
QY 149 LDVYERSLSNYPD-FQG-ARIITGQEGAGAWITITINLLKFSQKTWFSEIVETNNQ 206
| :
Db 684 ATFGALDGGSSLOVTEENERTHETNLNLRIGSVNHLISAYSILAGLNDAPDRSVYH 743
||||| :
QY 207 ETFGALDGGASTQVTEVPQQTIESPDNALQFRILYGRDYNVYHS-F-L--CYGKQDA- 261
| :
Db 744 LKRLPNVSKDLIEGLKEMKHPCLNSYN-GQYICSGC-ASSVQGGKSGSVSIKIVG 801
| :
QY 262 LMOKLAK-DI-QVANSNEI-LRDPCHPDKRVVNSDLYKPCPKRREMTLFOOFELQGG 318
| :
Db 802 APNNGEC--SALAK-NAP-C--AL-P-DG-YPRP-HGQFYVSGFVYRPFNLASAE-AS 850
| :
QY 319 IGNVQOCHOSTLELPNFSYCYSCAFNGITLPLPQGDGFAFAYPMKLNLTSEVYS 378
| :
Db 851 LDDVLEKRGREFCDKAMQVARTS-VSP-QPTEYOXCFAPIYVSLIREGL-YITDK-QI- 904
| :
QY 379 QEKYTEMMKRCQAQPMWEIKTSYGVKELYSCEGSCGTITLSLLQGHFTADSWEHII 438
| :
Db 905 TIGS--GSIT-WTIGVAL-L-E--SG-KALSTIGLAKSYFLSKINPINALISILSL 956
| :
QY 439 FIGIKQSDAGMTLGYMLNMIAPQDPLSTPLSHSYVPL-WVLESLVLTVAIIGLL 497
| :
Db 957 LL 958
| :
QY 498 IF 499

RESULT 6
ENTRY S48859 #type complete
TITLE nucleoside triphosphatase precursor, chromatin-associated -
ORGANISM garden pea
#formal_name Pisum sativum #common_name garden pea
DATE 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 29-Sep-1999
ACCESSIONS S65147; S48859
REFERENCE S65141
#authors Hsieh, H.L.; Tong, C.G.; Thomas, C.; Roux, S.J.
#journal Plant Mol. Biol. (1996) 30:135-147
#title light-modulated abundance of an mRNA encoding a calmodulin-regulated, chromatin-associated NTPase in pea.
#cross-references MUID:96197404
#accession S65147
##molecule-type mRNA
##residues 1-455 ##label HS2
##cross-references EMBL:Z32743; NID:9563611; PIDN:CAA83655.1;
PID:9563612

CLASSIFICATION #superfamily nucleoside triphosphatase chromatin-associated
KEYWORDS nucleus

S.Y.; Glaser, P.; Goffeau, A.; Golightly, E.J.; Grandi, G.
Guseppl, G.; Guy, B.T.; Haga, K.; Halech, J.; Harwood,
C.R.; Henaut, A.; Hildbert, H.; Holzapfel, S.; Hosono, S.;
Kusho, M.F.; Itaya, M.; Jones, L.; Joris, B.; Katamata, D.;
Kashara, Y.; Klauer-Blanchard, M.; Klein, C.; Kobayashi,
Y.; Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.;
Kurita, K.; Lapidus, A.; Lardinis, S.; Lauber, J.;
Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.;
Mausel, C.; Medigue, C.; Medina, N.; Mellido, R.P.; Mizuno,
M.; Moestl, D.; Nakai, S.; Noback, M.; Noone, D.; O'Reilly,
M.; Ogawa, K.; Ogiwara, A.; Outega, B.; Park, S.H.; Parro,
V.; Pohl, T.M.; Portetelle, D.; Potwolk, S.; Prescott,
A.M.; Presecan, E.; Pujić, P.; Purnelle, B.; Rapoport, G.;
Rey, M.; Reynolds, S.; Rieger, M.; Rivolta, C.; Rocha, E.;
Roche, B.; Rose, M.; Sadie, Y.; Sato, T.; Scanlon, E.;
Schleich, S.; Schroeter, R.; Scorrione, F.; Seliguch, U.;
Sekowska, A.; Seror, S.U.; Serrou, P.; Shin, B.S.; Soldo,
B.; Sorokin, A.; Tacconi, E.; Takagi, T.; Takahashi, H.;
Takamaru, K.; Takeuchi, M.; Tamokoshi, A.; Tanaka, T.;
Terstra, P.; Toironi, A.; Tosato, V.; Uchiyama, S.;
Vandenbol, M.; Vanlier, F.; Vassarotti, A.; Viari, A.;
Wamuntz, R.; Wedler, E.; Weller, H.; Welteneberger, T.;
Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto,
K.; Yata, K.; Yoshida, K.; Yoshikawa, H.F.; Zumstein, E.;
Yoshikawa, H.; Danchin, A.

Nature (1997) 390:249-256

The complete genome sequence of the Gram-positive bacterium
Bacillus subtilis.

#journal #title
#cross-references MOID:98044033
#accession G69998
#status preliminary; nucleic acid sequence not shown;
translation not shown

##molecule_type DNA
##residues 1-107 ##label KUN
##cross-references GB:259119; GB:LD009126; NID:g2635411;
PIDN:CAB14962.1; PID:e1185857; PID:g2635468

##experimental_source strain 168

GENETICS

#gene ytpP
CLASSIFICATION #superfamily thioredoxin; thioredoxin homology
SUMMARY #length 107 #molecular-weight 12747 #checks 1509

Query Match 3.0%; Score 116; DB 2; Length 107;
Best Local Similarity .34.8%; Pred. NO.1.42e-01;
Matches 16; Conservative 13; Mismatches 12; Indels 5; Gaps 4;

Dd 22 MESADWCP--DCRFVEPFLELANPEEF--YYVDNDKFIDTCAE 64
::: || :| | |||::: | :||| ::| :|
Oy 332 LNTSTCPYSQCAFNGLFLPLDLOGDFCAFNAFVV-KRLNLTSE 375

RESULT 13
ENTRY
TITLE S37786 #type complete
hypothetical protein YKL65c - yeast (*Saccharomyces cerevisiae*)
ALTERNATE_NAMES hypothetical protein YKL619
ORGANISM #formal_name *Saccharomyces cerevisiae*
DATE 03-May-1994 #sequence_revision 03-May-1994 #text_change 06-Feb-1998
ACCESSIONS S37786; S37995; S44563
REFERENCE
#authors Vandebol, M.; Bolle, P.; Dion, C.; Portetelle, D.; Hilger, F.
#submission submitted to the EMBL Data Library, September 1993
#description DNA sequencing of a 36.2 kb fragment located between the *FAS1* and *LAI4* loci of chromosome XI of *S. cerevisiae*.
#accession S37786
#molecule_type DNA
#residues 1-919 ##label VAN
##cross-references EMBL:Z26877; NID:g407482; PID:g407483
##experimental_source strain S28BC
REFERENCE S37976
Vandebol, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hilger,

```

#submission submitted to the Protein Sequence Database, March 1994
#accession S37995
##molecule_type DNA
##residues 1-919 ##label VA2
##cross-references EMBL:Z28165; NID:9486288; PID:9486289; MIPS:YK165c
##experimental_source strain S288C
REFERENCE S44563
#authors Yeandbol, M.; Bolle, P.A.; Dion, C.; Forcettelle, D.; Hliger, F.
#journal Yeast (1994) 10:35-40
#title DNA sequencing of a 36.2 kb fragment located between the PAST1 and LAP4 loci of chromosome XI of Saccharomyces cerevisiae
#accession S44563
##status nucleic acid sequence not shown; translation not shown
##molecule_type DNA
##residues 1-919 ##label VA3
##cross-references EMBL:Z26577; NID:9407482; PID:9407483
##experimental_source strain S288C
##note the nucleotide sequence was submitted to the EMBL Data Library, September 1993
GENETICS
#gene SGD:MCD4
##cross-references SGD:S0001648; MIPS:YK165c
#map_position 11L
SUMMARY #length 919 #molecular_weight 105692 #checksum 6222
Query Match 3.0%; Score 113; DB 2; Length 919;
Best Local Similarity 24.8%; Pred. No. 3,29e-01;
Matches 27; Conservative 35; Mismatches 37; Indels 10; Gaps 9;
Db 422 ENYLEQEAITTEELMOTTEGLHYLTYYWRIRIRIVVF-GF-VGVMFSTIFLKSTI 479
| | | : : : : : | : | : | : | : | : | : | : | : | : |
Qy 406 EKLSEYCFSTY-ILSLLOGYHF-TADSEIHIFIGKIGSDACWTL-GYMLNLITNMI 462
: : : : : : : : : : : : : : : : : : : : : : :
Db 480 LENYIDOKASPLSHAVGSGIGILNLTLYQSPNFVYLLFPLPYF 528
: : : : : : : : : : : : : : : : : : : : : : :
Qy 463 --PA-EQPLSTPLSHSTYVFLMYLFLVLF-TVAIIIGLIFPK-PSYFW 506
: : : : : : : : : : : : : : : : : : : : : : :
RESULT 14
ENTRY PXLNPD #type complete
TITLE H+-transporting ATPase (EC 3.6.1.35), plasma membrane -
Leishmania donovani
ALTERNATE_NAMES proton-transporting ATPase
ORGANISM Leishmania donovani
DATE 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 18-Jun-1999
ACCESSION A27124
REFERENCE A27124
#authors Meade, J.C.; Shaw, J.; Lemaster, S.; Gallagher, G.; Stringer, J.R.
#journal Mol. Cell. Biol. (1987) 7:3937-3946
#title Structure and expression of a tandem gene pair in Leishmania donovani that encodes a protein structurally homologous to eucaryotic cation-transporting ATPases.
#cross-references MUID:88122116
#accession A27124
##molecule_type DNA
##residues 1-974 ##label MEA
##cross-references GB:M17883; NID:g159291; PIDN:AAA29227.1; PID:g159294
##note The authors translated the codon AGA for residue 352 as Lys
CLASSIFICATION #superfamily Na+/K+-transporting ATPase alpha chain; ATPase
KEYWORDS nucleotide-binding domain homology
ATP; hydrogen ion transport; hydrolase; membrane protein;
phosphoprotein
FEATURE
495-670 #domain ATPase nucleotide-binding domain homology #label
ATN
351 #active_site Asp (aspartylphosphate intermediate)
#status predicted
#length 974 #molecular_weight 107476 #checksum 834
SUMMARY

```


(TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution Rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Fri May 5 08:47:07 2000; Maspar time 27.17 Seconds
Tabular output not generated. 560.539 Million cell updates/sec

Title: >US-09-374-586-1
Description: (1-510) from US09374586.pep
Perfect Score: 3806
Sequence: 1 MEDTRESNVKTRCSKNILAI.....VAIIIGLLIFHKPSYWKMY 510

Scoring table:
PAM 150
Gap 11

Searched: 82229 seqs, 29864866 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot38
1:swissprot

Statistics: Mean 51.996; Variance 95.108; scale 0.547

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	3806	100.0	510	CD39_HUMAN	VASCULAR ATP-DIPHOSPHO	0.00e+00
2	3056	80.3	510	CD39_MOUSE	VASCULAR ATP-DIPHOSPHO	0.00e+00
3	378	9.9	557	YB04_CAEEL	HYPOTHETICAL 63.1 KD P	1.19e-50
4	360	9.5	630	YE45_YEAST	HYPOTHETICAL 71.9 KD P	5.89e-47
5	324	8.5	485	Y44E_CAEEL	HYPOTHETICAL 54.3 KD P	1.18e-39
6	319	8.4	454	APY_SOLIU	ATPASE PRECURSOR (EC	1.18e-38
7	291	7.6	455	NPFA_PEA	NUCLEOSIDE-TRIPHOSPHAT	4.32e-33
8	265	7.0	518	GDAL1_YEAST	GUANOSINE-DIPHOSPHATAS	5.04e-28
9	125	3.3	592	NP2_TOKGO	NUCLEOSIDE-TRIPHOSPHAT	1.31e-03
10	127	3.3	628	NP2_TOKGO	NUCLEOSIDE-TRIPHOSPHAT	1.31e-03
11	125	3.3	628	NP2_TOKGO	NUCLEOSIDE-TRIPHOSPHAT	1.31e-03
12	117	3.1	528	UDBA_HUMAN	UDP-GLUCURONOSYLTRANSFERASE	1.77e-02
13	116	3.0	436	ACHX_ONCVO	ACETYLCHOLINE RECEPTOR	2.44e-02
14	113	3.0	919	YK05_YEAST	HYPOTHETICAL 105.7 KD	4.57e-02
15	114	3.0	974	ATXA_LEIDO	PROBABLE E1-E2 TYPE CA	6.99e-01
16	105	2.8	259	REFA2_KLEPN	O-ANTIGEN EXPORT SYSTEM	3.87e-01
17	107	2.8	276	NOSY_PSEST	MEMBRANE PROTEIN NOSY	3.87e-01
18	107	2.8	489	ACHB_XENIA	ACETYLCHOLINE RECEPTOR	2.87e-01
19	108	2.8	511	ACH5_CAEEL	ACETYLCHOLINE RECEPTOR	2.87e-01
20	106	2.8	521	ACHD_XENIA	ACETYLCHOLINE RECEPTOR	5.21e-01
21	106	2.8	573	COX1_MYCVU	PROBABLE CYTOCHROME C	2.87e-01
22	108	2.8	974	ATXB_LEIDO	PROBABLE E1-E2 TYPE CA	2.87e-01
23	102	2.7	450	TBA3_MAIZE	TUBULIN ALPHA-3 CHAIN	1.67e+00

ALIGNMENTS

RESULT	ID	CD39_HUMAN	STANDARD	PRT	510 AA.
AC	P49961				
DT	01-OCF-1996 (Rel. 34, Created)				
DT	15-DEC-1998 (Rel. 37, last sequence update)				
DE	VASCULAR ATP-DIPHOSPHOHYDROLASE (EC 3.6.1.5) (ATPase) (LYMPHOID CELL ACTIVATION ANTIGEN) (CD39 ANTIGEN).				
GN	CD39.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;				
OC	Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE; 95015846.				
RA	MALISZEWSKI C.R., DELESSESSE G.J.T., SCHOENBORN M.A., ARMSTRONG R.J., FANSLOW W.C., NAKAJIMA T., BAKER E., SUTHERLAND G.R., POINDEXTER K., BIRKS C., ALPERT A., FRIEND D., GIMPEL S.D., GAYLE R.B. III.				
RT	"The CD39 lymphoid cell activation antigen. Molecular cloning and structural characterization."				
RT	J. Immunol. 153(3):583-588(1994).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=UMBILICAL VEIN;				
RX	MEDLINE; 97149443.				
RA	ROBSON S.C., KACZMAREK E., SIEGEL J.B., CANDINAS D., KOZIAK K., MILLAN M., HANCOCK W.W., BACH F.H.				
RT	"Loss of ATP diaphosphohydrolase activity with endothelial cell activation."				
RT	J. Exp. Med. 185:153-163(1997).				
RN	[3]				
RP	FUNCTION.				
RA	KACZMAREK E., KOZIAK K., SEVIGNY J., SIEGEL J.B., ANRATHER J., BRADDOIN A.R., BACH F.H., ROBSON S.C.				
RT	"Identification and characterization of CD39/vascular ATP diaphosphohydrolase."				
RT	J. Biol. Chem. 271:33116-33122(1996).				
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).				
CC	-1- TISSUE SPECIFICITY: EXPRESSED PRIMARILY ON ACTIVATED LYMPHOID CELLS. ALSO EXPRESSED IN ENDOTHELIAL TISSUES.				
CC	-1- SIMILARITY: BELONGS TO THE GDAL / CD39 NPase FAMILY.				
CC	-1- DATABASE: NAME=PROV; NOTE=CD guide CD39 entry;				
CC	WWW=http://www.ncbi.nlm.nih.gov/prov/cd/cd39.htm"				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: S73813; AAB32152.1; -
DR EMBL: U87967; AAB47572.1; -
DR MIM: 601752; -
DR PROSITE: PS01238; GDA1_CD39_NTPASE. 1.

DR PFAM: PF01150; GDA1_CD39; 1.
KW Hydrolyase; Transmembrane; Antigen; Glycoprotein.
FT DOMAIN 1 16 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 17 38 POTENTIAL.
FT DOMAIN 39 478 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 479 499 POTENTIAL.
FT DOMAIN 500 510 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 73 73 POTENTIAL.
FT CARBOHYD 227 227 POTENTIAL.
FT CARBOHYD 292 292 POTENTIAL.
FT CARBOHYD 334 334 POTENTIAL.
FT CARBOHYD 371 371 POTENTIAL.
FT CARBOHYD 457 457 POTENTIAL.
SQ SEQUENCE 510 AA; 57964 MW; E403B5C9 CRC32;

Query Match 100.0%; Score 3806; DB 1; Length 510;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 MEDRESNVTFCFSKNILAILGFTSSIIAIVIALAVGLTQNKALPENVKYGIYLDAGSSHT 60
QY 1 MEDRESNVTFCFSKNILAILGFTSSIIAIVIALAVGLTQNKALPENVKYGIYLDAGSSHT 60
DB 61 SLIYKMPAEKENDGVVQVVECECRVKGPGISKFOKVEIGIYITDCMERAREVIIPRSQ 120
QY 61 SLIYKMPAEKENDGVVQVVECECRVKGPGISKFOKVEIGIYITDCMERAREVIIPRSQ 120
DB 61 HOETPVYLGATGMRLRMESEELADRVLDVVERLSNYPEDOGARITTOGEGAYGWI 180
QY 61 HOETPVYLGATGMRLRMESEELADRVLDVVERLSNYPEDOGARITTOGEGAYGWI 180
DB 121 HOETPVYLGATGMRLRMESEELADRVLDVVERLSNYPEDOGARITTOGEGAYGWI 180
QY 121 HOETPVYLGATGMRLRMESEELADRVLDVVERLSNYPEDOGARITTOGEGAYGWI 180
DB 181 TINYLLGKFSQKTRMFSIYVETNNQETFGALDVGASTQVTFVPONOTIESPDNALQFR 240
QY 181 TINYLLGKFSQKTRMFSIYVETNNQETFGALDVGASTQVTFVPONOTIESPDNALQFR 240
DB 181 TINYLLGKFSQKTRMFSIYVETNNQETFGALDVGASTQVTFVPONOTIESPDNALQFR 240
QY 181 TINYLLGKFSQKTRMFSIYVETNNQETFGALDVGASTQVTFVPONOTIESPDNALQFR 240
DB 241 LYGKYNYVTHSFLCYGKQDALMOKLADIOVASNEILRDPCHPGYKKVNVNSDLTKTP 300
QY 241 LYGKYNYVTHSFLCYGKQDALMOKLADIOVASNEILRDPCHPGYKKVNVNSDLTKTP 300
DB 241 LYGKYNYVTHSFLCYGKQDALMOKLADIOVASNEILRDPCHPGYKKVNVNSDLTKTP 300
QY 241 LYGKYNYVTHSFLCYGKQDALMOKLADIOVASNEILRDPCHPGYKKVNVNSDLTKTP 300
DB 301 CTKREMTLPFOQFPIQIGNTVOOCHOSTLEFNSTYCPYSCAFNGITFLPLOGDGFAP 360
QY 301 CTKREMTLPFOQFPIQIGNTVOOCHOSTLEFNSTYCPYSCAFNGITFLPLOGDGFAP 360
DB 361 SAFYVVMFLNTSEKVSQEKYTEMKKKFCQAPWEEIKTSYAGVKKKYSEYCFSGTYIL 420
QY 361 SAFYVVMFLNTSEKVSQEKYTEMKKKFCQAPWEEIKTSYAGVKKKYSEYCFSGTYIL 420
DB 421 SLLLOGYHTADSWEHIFIGIKIQSDAGWTIGYMLNTMIPAEOPUSTPLSHSTYVEL 480
QY 421 SLLLOGYHTADSWEHIFIGIKIQSDAGWTIGYMLNTMIPAEOPUSTPLSHSTYVEL 480
DB 481 MYLFSVLFTVAIIGLIFHKPSYFMKDMV 510
QY 481 MYLFSVLFTVAIIGLIFHKPSYFMKDMV 510

RESULT 2
ID CD39_MOUSE STANDARD; PRT; 510 AA.
AC P55772;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE VASCULAR ATP-DIPHOSPHODIOLASE (EC 3.6.1.5) (ATPDASE) (LYMPHOID CELL
DE ACTIVATION ANTIGEN) (CD39 ANTIGEN).
GN CD39.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 95015846.
RA MALISZEWSKI C.R., DELESPESE G.J.T., SCHOENBORN M.A., ARMITAGE R.J.,
RA FANSLAW W.C., NAKAJIMA T., BAKER E., SUTHERLAND G.R., POINDEXTER K.,
RA BIRKS C., ALPERT A., FRIEND D., GIMPEL S.D., GAYLE R.B. III;
RT "The CD39 lymphoid cell activation antigen. Molecular cloning and
RT structural characterization.";
RL J. Immunol. 153:3574-3583(1994).
CC -1 CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 PHOSPHATE.
CC -1 SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1 SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AF037366; AAB92259.1; -
DR MGD: MGI:102805; CD39.
DR PROSITE: PS01238; GDA1_CD39_NTPASE. 1.
DR PFAM: PF01150; GDA1_CD39; 1.
KW Hydrolyase; Transmembrane; Antigen; Glycoprotein.
FT DOMAIN 1 16 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 17 38 POTENTIAL.
FT DOMAIN 39 478 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 479 499 POTENTIAL.
FT DOMAIN 500 510 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 73 73 POTENTIAL.
FT CARBOHYD 226 226 POTENTIAL.
FT CARBOHYD 291 291 POTENTIAL.
FT CARBOHYD 333 333 POTENTIAL.
FT CARBOHYD 428 428 POTENTIAL.
FT CARBOHYD 457 457 POTENTIAL.
SQ SEQUENCE 510 AA; 57205 MW; 0570B8FE CRC32;

Query Match 80.3%; Score 3056; DB 1; Length 510;
Best Local Similarity 76.2%; Pred. No. 0.00e+00;
Matches 390; Conservative 68; Mismatches 50; Indels 4; Gaps 4;

DB 1 MEDIKSKVYKRCFSKNILAILGFTSSIIAIVIALAVGLTQNKALPENVKYGIYLDAGSSHT 60
QY 1 MEDIKSKVYKRCFSKNILAILGFTSSIIAIVIALAVGLTQNKALPENVKYGIYLDAGSSHT 60
DB 61 NLIYKMPAEKENDGVVQVVECECRVKGPGISKYAKTDEIGAYLAECMELSTELIPTSK 120
QY 61 NLIYKMPAEKENDGVVQVVECECRVKGPGISKYAKTDEIGAYLAECMELSTELIPTSK 120
DB 61 HOETPVYLGATGMRLRMESEELADRVLDVVERLSNYPEDOGARITTOGEGAYGWI 180
QY 61 HOETPVYLGATGMRLRMESEELADRVLDVVERLSNYPEDOGARITTOGEGAYGWI 180
DB 121 HOETPVYLGATGMRLRMESEELADRVLDVVERLSNYPEDOGARITTOGEGAYGWI 180
QY 121 HOETPVYLGATGMRLRMESEELADRVLDVVERLSNYPEDOGARITTOGEGAYGWI 180
DB 181 TINYLLGKFSQKTRMFSIYVETNNQETFGALDVGASTQVTFVPONOTIESPDNALQFR 240
QY 181 TINYLLGKFSQKTRMFSIYVETNNQETFGALDVGASTQVTFVPONOTIESPDNALQFR 240
DB 240 LYGEDYTVYTHSFLCYGKQDALMOKLADIOVSSGVLDKPCFNGYKVVVNSLYGTP 299
QY 241 LYGKYNYVTHSFLCYGKQDALMOKLADIOVASNEILRDPCHPGYKKVNVNSDLTKTP 300
DB 300 CTKREMTLPFOQFPIQIGNTVOOCHOSTLEFNSTYCPYSCAFNGITFLPLOGDGFAP 360
QY 301 CTKREMTLPFOQFPIQIGNTVOOCHOSTLEFNSTYCPYSCAFNGITFLPLOGDGFAP 360

```

Db 360 SAFYVADFFKAKNSVYSIOEKMTETKNCFSKSWETKTSTPSYKKEVSEYCFSGAY 419
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 361 SAFYVADFFKAKNSVYSIOEKMTETKNCFSKSWETKTSTPSYKKEVSEYCFSGAY 418
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 420 ILSLQGYNEFTDSSWEOIHPMKIKSDNAGWTGLYMLNTNMIPEOPLSPPLPHSTYI 478
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 419 ILSLQGYNEFTDSSWEOIHPMKIKSDNAGWTGLYMLNTNMIPEOPLSPPLPHSTYI 478
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 479 GLMYFLSLVAVATLGTFTSKSPSYKKEAV 510
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 479 FLMYFLSLVAVATLGTFTSKSPSYKKEAV 510
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |

RESULT 3
ID YBIA.CAEEL STANDARD; PRT; 557 AA.
AC Q21815;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHEICAL 63.1 KD PROTEIN R07E4.4 IN CHROMOSOME X.
GN R07E4.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditioidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA MILLER N.;
RL Submitted (NOV-1995) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
DR EMBL; U39652; AAA80403.1; -
DR PROSITE; R07E4.4; CE04819.
DR PFAM; PF01150; GDAL_CD39; 1.
KM Hypothetical protein, Transmembrane, Hydrolase.
FT TRANSMEM 7 27
FT TRANSMEM 495 515 POTENTIAL.
SQ SEQUENCE 557 AA; 63056 MW; 34CD5D10 CRC32;

Query Match 9.9%; Score 378; DB 1; Length 557;
Best Local Similarity 23.9%; Pred. No. 1.19e-50;
Matches 122; Conservative 134; Mismatches 217; Indels 37; Gaps 30;

Db 12 USAMIFPEVI-VFTYVVEAHTSPKVIADDERSYGVICDAGSTGTRLFVNMISTDSEL 70
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 18 LAIIGFSSIIAVIALGLQNKALPEN-VK-YGIYLDAGSSHTSLITYMPAKKENDT 75
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 71 IOIEPVYDNKPVAKKISPGISTGTYPQAQAEYLRPLMELAEKRHPPEYVPIFA 130
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 76 GVIVQVE-ECR-V--K-GPGISKFEVQKVEIGIYLTDCMERAREIPRSQHOETPVYIGA 130
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 131 TAGHRLIPDEVVLIGOKEAVLKNLKPITKSIOVLKEHRIITEGKKEGYSMTAVNYA 190
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 131 TAGHRLIPDEVVLIGOKEAVLKNLKPITKSIOVLKEHRIITEGKKEGYSMTAVNYA 190
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 191 LGRFNKATLDPGTPSAHAKOKTVGMIDMGASAOIAFELPDVDFESSIVENINLCGR 250
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 186 LGRFSQKTRM-FSTVPEYETNNQETFGALDGGASTQYTF-VPQNOTIES-P-DN-ALQFR 240
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 251 EDDSLFKFLFTVTFGLGVNEGIRKYEHLMLSLKIDKONGTVIIDDCLPLNLRKTVLEN 310
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 241 LYGR-K-DYNYVTHSFLCYGKQDALM--QKLAKDIQVASNIEILRDPGFHPGKRVVNSD 295
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 311 G-ENP-VARGGNMNTCSNEVKILLN-PESSEYCKA-EAAKCYFGAVPAISILSNTEM 366
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |

```

```

QY 296 LYKTPCTRFRFEMTLPLFOQFEILOGIGNYQCHQSLTELENTSYCPYSGQAFNGIFLPLQOG 355
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 367 -YG-FSEYWTSHDVLGIGO-YDAENIAKKTQOYCSKRWSTIQAESKKOLYPRADERL 423
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 356 DFGAFSAFYFVKM-FLNLTSEKVSQEKTEEMKKRCAQPW--E-EIKTS-YAGVKEKYL 409
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 424 RTQCFKSMWITSVLHDGS-VDKTHNKQSVSTIAGQVQVATLAMYTHMFFPLROSSR 482
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 410 SEYCFSGTYILSLQGYNEFTDSSWEOIHPMKIKSDNAGWTGLYMLNTNMIPEOPLS 469
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 483 NLIIVETHSSSESLMAPLFLSAVCFVL 512
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 470 TPLSHSTYVFLMYFLSLVAVATLGTFTSKSPSYKKEAV 510
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |

RESULT 4
ID YE45.YEAST STANDARD; PRT; 630 AA.
AC P40009;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHEICAL 71.9 KD PROTEIN IN PM140-PAC2 INTERGENIC REGION.
GN YER005W.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;
RA DIETRICH F.S., MULLIGAN J.T., HENNESSEY K.M., ALLEN E., ARAUJO R.,
RA AVILES E., BERNI A., BRENNAN T., CARPENTER J., CHEN E., CHERRY J.M.,
RA CHUNG E., DUNCAN M., GUZMAN E., HARTZELL G., HONICKER-SMITH S.,
RA HYMAN R., KAYSER A., KOMP C., LASHKARI D., LEW H., LIN D.,
RA MUSEDLE D., NAKAHARA K., NAMATH A., NORGEN R., OEFNER P., OH C.,
RA PETERL F.X., ROBERTS D., SEHL P., SCHRAMM S., SHOGEN T., SMITH V.,
RA TAYLOR P., WEI Y., YELTON M., BONSTEIN D., DAVIS R.W.;
RL Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.
-----
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
DR EMBL; U18778; AAB64538.1; -
DR PROSITE; PS01238; GDAL_CD39_NTPASE; 1.
DR PFAM; PF01150; GDAL_CD39; 1.
KM Hypothetical protein, Transmembrane, Hydrolase.
FT TRANSMEM 501 517 POTENTIAL.
SQ SEQUENCE 630 AA; 71851 MW; 68CD0D15 CRC32;

Query Match 9.5%; Score 360; DB 1; Length 630;
Best Local Similarity 28.4%; Pred. No. 5.89e-47;
Matches 106; Conservative 93; Mismatches 137; Indels 37; Gaps 28;

Db 52 IHQKDWTFKLNPGISPEKKRQDAYKSHIKPLDFAKNIIIPESWMSQCFPIQATAGMR 111
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 78 VHQVECKRVK-GPGISKFEVQKVEIGIYLTDCMERAREIPRSQHOETPVYIGA 135
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 112 LLPDIOSSIIIDGLCGIKHP-AEFLVEDCSAQIOIVIDGETGLGYGLNLYLGHFND- 169
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 136 LIRNESEE-LADRYLDVYERSLSNYPF-DFQGA-RITGQEGAYGWTIYLLGKRSQK 192
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 170 ---YN--P-EVSDHFTFGMDMGASQTOIAPAPDSGEIARHARDIAFIPLRSVNGDLQK 223
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
QY 193 TRMSIVPEYETNNQETFGALDGGASTQYTFVPQNOT-IESP-DN-ALQF-R-LYGR-D- 245
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 224 MDVFSVTHLFGANQARRRYLAOLINTLPENT-ND-YENDQFSRR-NLNDMCDRGSSTD 280
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |

```

[illegible]

RESULT	5	STANDARD:	PRT:	485 AA.
ID	YVAE.CAEEL			
AC	018411;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	HYPOTHELTICAL 54.3 KD PROTEIN C33H5.14 IN CHROMOSOME IV.			
GN	C33H5.14.			
OS	Caenorhabditis elegans.			
OC	Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;			
OC	Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-BRISTOL N2.			
RA	BRADSHAW H., STEELYES L.;			
RL	Submitted (DEC-1995) to the EMBL/Genbank/DBJ databases.			
CC	-1 SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; U41007; AAA82272.1; -.			
DR	WORMPEP; C33H5.14; CE04A157.			
DR	PROSITE; PS01238; GDAL_CD39__NTPASE; 1.			
KM	PF01150; GDAL_CD39; 1.			
KM	Hypothetical protein; Transmembrane; Hydrolase.			
FT	TRANSMEM 439 459 POTENTIAL.			
EQ	SEQUENCE 485 AA; 54309 MW; 00659E2B CRC32;			

Query Match	8.5%;	Score 324;	DB 1;	Length 485;
Best Local Similarity	24.9%;	Pred. No. 1.18e-39;		
Matches	116;	Conservative 113;	Mismatches 196;	Indels 40;
			Gaps 34	

D	b	21	NNIKKGVACDSSGSTRFEVYTLPLPSGLNIDITLHESPPVKATYPGLSFGDPQ	80
O	y	45	ENNAKKGYLDAGSSHTSLYIKW-P-AEKENDTG-VNHQVEBCAVK-GPGISKVQAYNE	1009
D	b	81	VVEYLPLRFAEHHPIPEQLGETDLLIFATGRLLPEAKODAI IKNLGNLKSTVALR	1409
O	y	101	IGIYLTDCMERARERIVPSCHOETPVYLGATGRMLTMESEELADVLIVERLSLNP	1609
D	b	141	VSDSNIRIIDAMEGISWIAVNVTILRPEDKENS-KYGMIDMGASVOJAFELANE-KE	1989
O	y	161	FDFOGARITIQEGCAIGMTINTLLGGFSOKTFWESTVPEYETNNQETFGALDGAStQ	2209
D	b	199	-SYNGCN-VYEINGISJETINEDY-K-YKIYSTPLGAGANGCL-KKYENSL-VKSNGS-N	2519
O	y	221	VTFWPOQOTIESPDNALQFRL-YEKDVNVVYHSPLCYGKDQALMQKLADIQVANSIELR	2799
D	b	252	DSCSPRGKNLI--GE-FTVANGTEMBDYCLA-QVYSSL-TGD-KA-QPSC-P--NPL-CF	2999
O	y	280	DPCHNPGRKKVANVSCLKTPCTLRPFETTLRFPOOFLQGIQANQQOCHQSILELFNSYCP	3399

```

Db 300 LRNVIAPSVNLSTVOL-YG-SSEWYTTSSNGSSGGEHYQ-KFIDEXRYKKQKMDNIOD 356
OY 340 YSQACNFCNIFLPLOGDGGAFAFYFNKKFLNTLNTSEKYSQKVEYEMKKKCAQDPWEEI-- 397
Db 357 GFKRNEFPNADIERLTGTCNCFKAAMWTSLVHDGFN-VDKTKLFSQVYLKIGEEKQNALGA 415
OY 388 --KTS-YGVGVEKYLYSEYCFSGTITLTLSDLOGHYHPTADSWENHIFIKIGSDGKWTGTY 454
Db 416 MLYHSKDKLFNLLEQLEVAQSTQOISNFF-SFF-VILIIVLAVAL 458
OY 455 ML-NLTNM-IRAEQPLSTPLS-HSYIVYVLAMLFESVYLTAVALISL 496

```

ID	RESULT
6	
APY_SOUTU	STANDARD:
PRT;	454 AA.
RX R80595; 043164:	
DT 01-OCT-1996 (Rel. 34, Created)	
DT 01-NOV-1997 (Rel. 35, Last sequence update)	
DE APYRASE PRECURSOR (EC 3.6.1.5) (ATP-DIPHOSPHATASE) (ADENOSINE DIPHOSPHATASE) (ADPase) (ATP-DIPHOSPHOHYDROLASE).	
GN RPOPL.	
OS Solanum tuberosum (Potato).	
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; eudicotyledons; Spermatophyta; Magnoliopsida; euasterids I; Solanales; Solanaceae; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; solanum.	
[1]	
RP SEQUENCE FROM N.A., AND SEQUENCE OF 59-160; 236-253 AND 332-345.	
RN TISSUE-TUBER.	
RC MEDLINE; 96158985.	
RK HANDA M., GUIDOTTI G. ;	
RA "Purification and cloning of a soluble ATP-diphosphohydrolase [apyrase] from potato tubers (solanum tuberosum)." ;	
RT Biochem. Biophys. Res. Commun. 218:916-923(1996).	
RL [2]	
RP SEQUENCE OF 42-54; 68-95 AND 236-253.	
RP STRAIN-CV. DESIREE.	
RX MEDLINE; 96955615.	
RA VASCONCELOS E.G.; FERREIRA S.T.; DE CARVALHO T.M.U.; DE SOUZA W.; KETTLUNG A.M.; MANCILLA M.; VALENZUELA M.A.; VERJOVSKI-ALMEIDA S.;	
RT "Partial purification and immunohistochemical localization of ATP diphosphohydrolase from Schistosoma mansoni. Immunological cross-reactivities with potato apyrase and Toxoplasma gondii nucleoside triphosphate hydrolase." ;	
RT J. Biol. Chem. 271:22139-.22145(1996).	
CC -1 FUNCTION: CATALYZES THE HYDROLYSIS OF PHOSPHOANHYDRIDE BONDS OF NUCLEOSIDE TRI- AND DI-PHOSPHATES.	
CC -1 CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 PHOSPHATE.	
CC -1 COFACTOR: CALCIUM.	
CC -1 SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED (PROBABLE).	
CC -1 PWM: THE N-TERMINUS IS BLOCKED.	
CC -1 SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.	
CC CC	
CC This SWISS-PROT entry is copyright. It is produced through a collab between the Swiss Institute of Bioinformatics and the EMBL outstc the European Bioinformatics Institute. There are no restrictions cc use by non-profit institutions as long as its content is in modified and this statement is not removed. Usage by and for commtified entities requires a license agreement (See http://www.isb.ch/an or send an email to llicense@isb-sib.ch).	
CC -----	
DR EMBL: U58597; AAC02720.1; -	
DR PROSITE; PS01238; GDAL_CD39_NTPASE. 1.	
DR Pfam: PF01150; GDAL_CD39; 1.	
KW Hydrolyase; Transmembrane; Calcium; Signal.	
FT SIGNAL 1 30 POTENTIAL.	
FT CHAIN 31 454 APYRASE.	
FT TRANSMEM 426 446 POTENTIAL.	
FT CARBOXYD 151 151 POTENTIAL.	
FT CARBOXYD 262 262 POTENTIAL.	
SQ SEQUENCE 454 AA; 50041 MW; D6FAEA89 CRC32;	

```

Query Match      8.4%; Score 319; DB 1; Length 454;
Best Local Similarity 28.3%; Pred. No. 1,18e-38;
Matches 69; Conservative 66; Mismatches 99; Indels 10; Gaps 9;

Db 12 ILAIFLVPLSLSKNVNAOIPLRHLLSHSEHYAVIFDAGSTGSRVHYFR-DEKTLGL 70
| | | | | : : : : : : : : : : : : : : : : : : : : : : : : : :
| | | | | : : : : : : : : : : : : : : : : : : : : : : : : : :
17 ILAI-LGFS-SIIAVIALAVGLTQNKALPENKYGIVLDAGSSHTSLYIKYKPAEKEND 74

Db 71 LPIGNINIEYMAIEPGLSSIAEDPKKAANSLEPLLDGAEVYQEOLOSEPLELGATAGL 130
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
75 TGVVHVEECRVKPGISKVFQKVNELGILYITDCMEARARVPRSOQETPPVYLGAATAGM 134

Db 131 RMKGDAEAKILAVRNLY-KNOSTFHSKDWYITLDGTQGSYMAAIVYLGNLG-KD 188
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
135 RLRLMSEEF-LADRVLDIVERSLSNYPFDQGARITIGQEGAGWITITILGKRSQKT 193

Db 189 -YKSTT-ATIDLG-GSVQMAVAISNEQPAKAPQNEDEGPYQKHLMSKDYNLVYHSY 244
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
194 RMFSIYVETNNQETFGALDLGASITQVTFVPQNOTIESPDNLQFRLYGKDYNVYTHSF 253

Db 245 LNYG 248
| |
QY 254 LCYG 257

RESULT 7
ID NTPA-PEA STANDARD; PRT; 455 AA.
AC P52914;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE NUCLEOSIDE-TRIPHOSPHATASE (EC 3.6.1.15) (NUCLEOSIDE TRIPHOSPHATE
DE PHOSPHOHYDROLASE) (NTPASE).
OS Pisum sativum (Garden pea).
OC Caryophyta; Viridiplantae; Streptophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;
OC Pisum.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV, ALASKA; TISSUE=PLUMULE;
RX MEDLINE; 96197404.
RA HSIEH H., TONG C.G., THOMAS C., ROUX S.J.;
RT "Light-modulated abundance of an mRNA encoding a
calmodulin-regulated, chromatin-associated NTPase in pea.";
RL Plant Mol. Biol. 30:135-147(1996).
CC -1- FUNCTION: MIGHT BE INVOLVED IN RNA TRANSPORT OUT OF NUCLEI.
CC -1- CATALYTIC ACTIVITY: NTP + H(2)O = NDP + PHOSPHATE.
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEUS.
CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: 232743; CAAB3655.1; -
CC PROSITE: PS01238; GDAL_CD39_NTPASE; 1.
CC PRAM: PF01150; GDAL_CD39; 1.
CC K01 Hydrolyase; Nuclear protein.
CC SEQUENCE 455 AA; 50072 MW; 38181E4C CRC32;

Query Match      7.6%; Score 291; DB 1; Length 455;
Best Local Similarity 28.2%; Pred. No. 4,32e-33;
Matches 72; Conservative 67; Mismatches 98; Indels 18; Gaps 16;

Db 6 KLITFLFSMPAITSQYIGNLLTSKRIFLKOEEISSVAVVDAGSTGSRHYFR-NQ 64
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
16 NIIAIIIGFS-SIIAVIALAVGL-TQNKAL-PENVK-YGIVLDAGSSHTSLYIKWPAE 70

```

```

Db 65 NIDLHITGKGEVYNNKITTPGLSSYANNPEQAKSLIPLRQAEVDVDDIQLQKPPVRLGA 124
: : : : : | | : : : : : : : : : : : : : : : : : : : : :
QY 71 KENDTGVVHVEECRVKPGISKVFQKVNELGILYITDCMEARARVPRSOQETPPVYLGA 130

Db 125 TAGILRL-NGDA-SEKILQSVRDMLSNRSTFNVPQDPAVSIIDTQGSYLMYVNLALG 181
| | | | | : : : : : : : : : : : : : : : : : : : : : : : :
131 TAGKRLRLMSEELADRVLDIVERSLSN-YPFDQ-GA-RITIGQEGAGWITITINVLIG 187

Db 182 NLGKRYTKYGVYDLDGGSVQV-AYAVSKTKAK-N-APKVAQDDP-YIKKVVLLKQIPY 236
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
188 KFSQK-TRMFSIYVETNNQETFGALDLGASITQVTFVPQNOTIESPDNLQFRLYGKDY 246

Db 237 DLYVHSLYHFGREAS 251
: : | | | : | : : : :
QY 247 NVYTHSFICYGKDOA 261

RESULT 8
ID GDAL YEAST STANDARD; PRT; 518 AA.
AC P32621;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE GUANOSINE-DIPHOSPHATASE (EC 3.6.1.42) (GDPASE).
GN GDAL OR YEL042W OR SYGP-ORF16.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomyces.
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN-G2-9;
RX MEDLINE; 93308137.
RA ABELION C., YANAGISAWA K., MANDON E.C., HAUSLER A., MOREMEN K.,
RA HIRSCHBERG C.B., ROBBINS P.W.;
RT "Guanosine diphosphatase is required for protein and sphingolipid
RT glycosylation in the Golgi lumen of Saccharomyces cerevisiae.";
RL J. Cell Biol. 122:307-323(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA MULLIGAN J.T., DIETRICH F.S., HENNESSEY K.M., SEHL P., KOMP C.,
RA WEI Y., TAYLOR P., NAKAHARA K., ROBERTS D., DAVIS R.W.;
RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;
RA DIETRICH F.S., MULLIGAN J.T., HENNESSEY K.M., ALLEN E., ARAUJO R.,
RA AVILES E., BERNARD A., BRENNAN T., CARPENTER J., CHEN E., CHERRY J.M.,
RA CHUNG E., DUNCAN M., GUZMAN E., HARTZELL G., HUNICKER-SMITH S.,
RA HYMAN R., KAYSER A., KOMP C., LASHKARI D., LEM H., LIN D.,
RA MOSEDALE D., NAKAHARA K., NAWATH A., NORGEN R., OERNER P., OH C.,
RA PATEL F.X., ROBERTS D., SEHL P., SCHRAMM S., SHOGREN T., SMITH V.,
RA TAYLOR P., WEI Y., YELTON M., BOTSTEIN M., DAVIS R.W.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: AFTER TRANSFER OF SUGARS TO ENOGENOUS MACROMOLECULAR
ACCEPTORS, THE ENZYME CONVERTS NUCLEOSIDE DIPHOSPHATES TO
NUCLEOSIDE MONOPHOSPHATES WHICH IN TURN EXIT THE GOLGI LUMEN IN
A COUPLED ANTIPORON REACTION, ALLOWING ENTRY OF ADDITIONAL
NUCLEOTIDE SUGAR FROM THE CYTOSOL.
CC -1- CATALYTIC ACTIVITY: GDP + H(2)O = GMP + ORTHOPHOSPHATE.
CC -1- PATHWAY: GLYCOSYLATION.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI.
CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L19560; AAA34656.1; -
CC DR EMBL: U18779; AAB65000.1; -

```



```

RT   Toxoplasma gondii".
RL   J. Biol. Chem. 270:11391-11397(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN-RH:
RX   MEDLINE: 95050750.
RA   BERNHEDS D., PECK K.R., AFIFI M.A., BECKERS C.J.M., JOINER K.A.;
RT   "randomly repeated genes encode nucleoside triphosphate hydrolase
RT   isoforms secreted into the parasitophorous vacuole of Toxoplasma
RT   gondii."
RL   J. Biol. Chem. 269:29252-29260(1994).
CC   -1- FUNCTION: MAY PERFORM AN IMPORTANT PROCESSING STEP IN THE
CC   CONVERSION OF HIGH ENERGY NUCLEOTIDES PRIOR TO UPTAKE BY THE
CC   PARASITE AND MAY CONTRIBUTE TO INTRACELLULAR SURVIVAL AND
CC   VIRULENCE. NTPASE-I HAS A SPECIFIC ACTIVITY 4.5-FOLD HIGHER THAN
CC   NTPASE-II IN HYDROLYSIS OF AMP. THE PRIMARY DIFFERENCE BETWEEN
CC   THESE ISOZYMS LIES IN THEIR ABILITY TO HYDROLYZE NUCLEOSIDE
CC   TRIPHOSPHATE VERSUS DIPHOSPHATE SUBSTRATES. WHILE NTPASE-II
CC   HYDROLYZES ATP TO ADP AND ADP TO AMP AT ALMOST THE SAME RATE,
CC   NTPASE-I HYDROLYZES ADP TO AMP AT A MUCH SLOWER RATE (0.7% OF THE
CC   RATE FOR ATP).
CC   -1- CATALYTIC ACTIVITY: NTP + H(2)O = NDP + PHOSPHATE.
CC   -1- SUBUNIT: HOMOTETRAMER.
CC   -1- SUBCELLULAR LOCATION: SECRETED; FOUND IN HOST CELL PARASITOPHOUS
CC   VACUOLE.
CC   -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.
CC   CC
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@lsb-sib.ch).
CC   -----
CC   EMBL, L39078; AAA89203.1; -.
CC   DR   EMBL, U96965; AAC80188.1; -.
CC   DR   PROSITE; PS01238; GDAL_CD39_NTPASE; 1.
CC   DR   PFMW; PF01150; GDAL_CD39; 1.
CC   KW   Hydrolase; Multigene family; Signal.
CC   FT   SIGNAL 1 25
CC   FT   CHAIN 26 628 NUCLEOSIDE-TRIPHOSPHATASE I.
CC   FT   CARBOHYD 432 432 POTENTIAL.
CC   FT   FT
CC   SQ   SEQUENCE 628 AA; 69159 MW; DA2A1577 CRC32;
CC
Query Match 3.3%; Score 125; DB 1; Length 628;
Best Local Similarity 32.8%; Pred. No. 1,31e-03;
Matches 22; Conservative 20; Mismatches 23; Indels 2; Gaps 2;
Db 228 TRPITGAEGLFAFTITLWHLSTRLEGDPAKMDIEGYKQKCRNDLAGVVEGASQAQIV 287
      :|||||::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY  166 ARIITGEGAGATWITTYLLGKFSQKRWMSIYVEIINN-QETP-GALDGGASTGYTF 223
      ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 288 PLOECTV 294
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY  224 VPONQTI 230
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

RX METLINE; 93326164.
RA JIN C.-J., MINERS J.O., LILLWHITE K.J., MACKENZIE P.I.;
RT "cDNA cloning and expression of two new members of the human liver
RT UDP-glucuronosyltransferase 2B subfamily.";
RL Biochem. Biophys. Res. Commun. 194:496-503(1993).
CC -1- FUNCTION: UDPGT IS OF MAJOR IMPORTANCE IN THE CONJUGATION AND
CC SUBSEQUENT ELIMINATION OF POTENTIALLY TOXIC XENOBIOTICS AND
CC ENDOGENOUS COMPOUNDS.
CC -1- CATALYTIC ACTIVITY: UDP-GLUCURONATE + ACCEPTOR = UDP + ACCEPTOR
CC BETA-D-GLUCORONOSIDE.
CC -1- SUBCELLULAR LOCATION: MICROSOMAL.
CC -1- SIMILARITY: BELONGS TO THE UDP-GLUCOSYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; X63359; CAA44961.1; -.
CC PIR; J06620; J06620.
CC MIM; 600070; -.
CC DR PROSITE; PS00375; UDPGT; 1.
CC PFM; PF00201; UDPGT; 1.
CC KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane; Signal;
CC Multigene family; Microsome.
CC FT SIGNAL 1 23 BY SIMILARITY.
CC FT CHAIN 42 528 UDP-GLUCURONOSYLTRANSFERASE 2B10.
CC FT TRANSMM 492 512 POTENTIAL.
CC FT CARBOHYD 66 66 POTENTIAL.
CC FT CARBOHYD 314 314 POTENTIAL.
CC FT CARBOHYD 481 481 POTENTIAL.
CC SQ SEQUENCE 528 AA; 60774 MW; 63F280A6 CRC32;

Query Match 3.1%; Score 117; DB 1; Length 528;
Best Local Similarity 24.7%; Pred. No.1,77e-02;
Matches 21; Conservatve 24; Mismatches 36; Indels 4; Gaps 4;

Db 133 KILMKIOESRPDIYFADADYLPCGELIAELFNIPFY-YSH-SFSPGYSPERHSGGF-TFP 189
OY 303 KRFEMTLPRQGEIIGICNGYQCHOSILELFTMTSCPYQCAFNGLF-LPPLGDPGAFS 361
Db 190 PSYVPVYMSKLSDDMTMERVRNML 214
OY 362 AEFYVWKFLNTSEKVSQEKVTEMM 386

RESULT 13
ID ACHX_ONCVO STANDARD; PRT; 436 AA.
AC P54247;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE ACETYLCHOLINE RECEPTOR PROTEIN, NON-ALPHA CHAIN (FRAGMENT).
OS Onchocerca volvulus.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Spirurida; Spirurida;
OC Filarioidea; Onchocercidae; Onchocerca.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94299155.
RA AJUH P.-M., EGWANG T.G.;
RT "Cloning of a cDNA encoding a putative nicotinic acetylcholine
RT receptor subunit of the human filarial parasite Onchocerca
RT volvulus.";
RL Gene 144:127-129(1994).
CC -1- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN
CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND
CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA
CC MEMBRANE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
CC EMBL, L20465; AAA21823.1; -.
CC DR EMBL, L12543; AAA29415.1; -.
CC DR PROSITE, PS00236; NEUROTR_ION_CHANNEL, 1.
CC DR PFAM: PF00065; neur_chan, 1.
CC Postsynaptic membrane; Ionic channel; Glycoprotein; Transmembrane;
CC Multigene family.
CC
CC FT NON_TER 1 1 EXTRACELLULAR (POTENTIAL).
CC FT DOMAIN <1 195 POTENTIAL.
CC FT TRANSMEM 196 219 POTENTIAL.
CC FT TRANSMEM 227 245 POTENTIAL.
CC FT TRANSMEM 261 280 POTENTIAL.
CC FT DOMAIN 281 404 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 405 423 POTENTIAL.
CC FT DISULFID 89 103 BY SIMILARITY.
CC FT CARBOHYD 62 62 POTENTIAL.
CC FT CARBOHYD 140 140 POTENTIAL.
CC SO SEQUENCE 436 AA; 51340 MW; 13E781FD CRC32;
CC
CC Query Match 3.0%; Score 116; DB 1; Length 436;
CC Best Local Similarity 38.5%; Pred. No. 2,44e-02;
CC Matches 20; Conservative 17; Mismatches 9; Indels 6; Gaps 6;
CC
CC Db 241 YLL-LKDIIPAT-STALPL-FEKKYLLFTMIMWSLSVL-VYVSLNLRHRRS 288
CC Oy 454 YMLNLTNMIIPAPLSTPLSHSTY-FLMWLSFLTVAIIGL-LIPHKPS 503
CC
CC RESULT_14
CC ID YK05_YEAST STANDARD; PRT; 919 AA.
CC AC P36051.
CC DT 01-JUN-1994 (Rel. 29, Created)
CC DT 01-JUN-1994 (Rel. 29, Last sequence update)
CC DT 01-OCT-1994 (Rel. 30, Last annotation update)
CC DE HYPOTHETICAL_105.7 KD PROTEIN IN TPKR-PIRL INTERGENIC REGION.
CC GN YKL165C OR YKL619.
CC OS Saccharomyces cerevisiae (Baker's yeast).
CC OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
CC OC Saccharomycetaceae; Saccharomycetes.
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN-S288C;
CC RX MEDLINE: 94378720.
CC RA VANDENBOL M., BOLE P.-A., DION C., PORTETELLE D., HILGER F.;
CC RT "DNA sequencing of a 36.2 kb fragment located between the FAS1 and
CC RT LAP loci of chromosome XI of Saccharomyces cerevisiae.";
CC YL 10:535-540(1994).
CC
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
CC DR EMBL, Z26877; CA81489.1; -.
CC DR EMBL, Z28165; CA82007.1; -.
CC DR PIR, S37786; S37786.
CC DR PIR, S44563; S44563.
CC KW Hypothetical protein.
CC SO SEQUENCE 919 AA; 105693 MW; 73C93BED CRC32;
CC
CC Query Match 3.0%; Score 113; DB 1; Length 919;
CC Best Local Similarity 24.8%; Pred. No. 6,24e-02;

```


(PL)

```

FT DOMAIN 38 478 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 499 499 POTENTIAL.
FT DOMAIN 500 511 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 73 73 POTENTIAL.
FT CARBOHYD 226 226 POTENTIAL.
FT CARBOHYD 291 291 POTENTIAL.
FT CARBOHYD 333 333 POTENTIAL.
FT CARBOHYD 374 374 POTENTIAL.
FT CARBOHYD 429 429 POTENTIAL.
FT CARBOHYD 458 458 POTENTIAL.
SQ SEQUENCE 511 AA; 57408 MW; E59DABG9 CRC32;

Query Match 79.8%; Score 3038; DB 11; Length 511;
Best Local Similarity 74.2%; Pred. No. 0.00e+00;
Matches 380; Conservative 73; Mismatches 56; Indels 3; Gaps 3;

Db 1 MEDTRSVKRCCKNIIILIGFSSVLAIVLAVGLTHNKPENVKYGIYLDAGSSHT 60
1 MEDTRSVKRCCKNIIILIGFSSVLAIVLAVGLTHNKPENVKYGIYLDAGSSHT 60
61 NLIYIKWPAEKENDGVVQLLEECOVKPGISKYAKOTDEIAVLAECMKMTERIPAK 120
61 SLIYIKWPAEKENDGVVQLLEECOVKPGISKYAKOTDEIAVLAECMKMTERIPAK 120
61 SLIYIKWPAEKENDGVVQLLEECOVKPGISKYAKOTDEIAVLAECMKMTERIPAK 120

Db 121 QHCPVYLGATAGMRLRMESKQADEVLAAYSRSLKSYPFDFOGAKITGOEGAYGWI 180
121 QHCPVYLGATAGMRLRMESKQADEVLAAYSRSLKSYPFDFOGAKITGOEGAYGWI 180
121 HQEPVYLGATAGMRLRMESKQADEVLAAYSRSLKSYPFDFOGAKITGOEGAYGWI 180

Db 181 TINYLLGKFTQBSWLANIS-DQKQATFGALDGGSSQTQVFPVPLNQTLEAPETSLOFR 239
181 TINYLLGKFTQBSWLANIS-DQKQATFGALDGGSSQTQVFPVPLNQTLEAPETSLOFR 239
181 TINYLLGKFTQBSWLANIS-DQKQATFGALDGGSSQTQVFPVPLNQTLEAPETSLOFR 240

Db 240 LYGDYVYVTSFLCYGQDQALMOKLADIOVASSGILKDPFPGYKVVVNSLQYPR 299
240 LYGDYVYVTSFLCYGQDQALMOKLADIOVASSGILKDPFPGYKVVVNSLQYPR 299
241 LYGDYVYVTSFLCYGQDQALMOKLADIOVASSGILKDPFPGYKVVVNSLQYPR 300

Db 300 CTKREKPLPENOQVOGTGDEYEOCHQSLIKFPNNSHCPSYSCAFNGVFLPPLQSGFAP 359
300 CTKREKPLPENOQVOGTGDEYEOCHQSLIKFPNNSHCPSYSCAFNGVFLPPLQSGFAP 359
301 CTKREKPLPENOQVOGTGDEYEOCHQSLIKFPNNSHCPSYSCAFNGVFLPPLQSGFAP 360

Db 360 SAFYVMDFFPKMANDSVSQEKMETIKNFKSKPWEVKASYPIVYKEKYLSEYCSGTY 419
360 SAFYVMDFFPKMANDSVSQEKMETIKNFKSKPWEVKASYPIVYKEKYLSEYCSGTY 419
361 SAFYVMDFFPKMANDSVSQEKMETIKNFKSKPWEVKASYPIVYKEKYLSEYCSGTY 418

Db 420 ILSLLQGYNFTGSTMDOIHFMWIKIDSNAGWTGLMNLTMIPAEOPLSPPLPSTI 479
420 ILSLLQGYNFTGSTMDOIHFMWIKIDSNAGWTGLMNLTMIPAEOPLSPPLPSTI 479
419 ILSLLQGYNFTGSTMDOIHFMWIKIDSNAGWTGLMNLTMIPAEOPLSPPLPSTI 478

Db 480 SLAVLFSLVLAAMVTGLTFESKPSYFWKNAV 511
480 SLAVLFSLVLAAMVTGLTFESKPSYFWKNAV 511
479 FLAVLFSLVLAAMVTGLTFESKPSYFWKNAV 510

RESULT 2 PRELIMINARY; PRT; 513 AA.
ID 018956;
AC 018956;
DT 01-JAN-1998 (Tremblrel. 05, Last created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
DE VASCULAR ATP-DIPHOSPHOHYDROLASE (EC 3.6.1.5) (LYMPHOID CELL ACTIVATION
DE ANTIGEN) (CD39 ANTIGEN) (ATPASE) (ECOTO-ADPASE) (ECOTO-
DE ADPASE) (NUCLEOTIDE PHOSPHOHYDROLASE) (ATP PYROPHOSPHOHYDROLASE).
OS Bos taurus (Bovinae).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae;
OC Bovinae; Bos.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-AORTIC ENDOTHELIUM;
RA CHANG A.S., GARCIA R.L., CHANG S.M., SCHILLING W.P.;
RL Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: HYDROLYSES EXTRACELLULAR ATP AND ADP TO AMP. COULD
INHIBIT PLATELET AGGREGATION IN RESPONSE TO ADP, COLLAGEN, AND

```

```

CC THROMBIN (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 PHOSPHATE.
CC -1- COFACTOR: CALCIUM ION-AND MAGNESIUM ION-DEPENDENT ACTIVITY (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
DR EMBL; AF005940; AAB62382.1; -
DR PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
DR PFM; PF01150; GDA1_CD39; 1.
KM Hydrolyase; Transmembrane; Antigen; Glycoprotein.
FT DOMAIN 1 16 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 17 37 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 38 481 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 482 502 POTENTIAL.
FT DOMAIN 503 513 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 73 73 POTENTIAL.
FT CARBOHYD 227 227 POTENTIAL.
FT CARBOHYD 245 245 POTENTIAL.
FT CARBOHYD 307 307 POTENTIAL.
FT CARBOHYD 336 336 POTENTIAL.
FT CARBOHYD 373 373 POTENTIAL.
FT CARBOHYD 460 460 POTENTIAL.
SQ SEQUENCE 513 AA; 58113 MW; 17735933 CRC32;

Query Match 74.8%; Score 2845; DB 6; Length 513;
Best Local Similarity 70.8%; Pred. No. 0.00e+00;
Matches 364; Conservative 83; Mismatches 62; Indels 5; Gaps 4;

Db 1 MEDRESELKVCCKNIIILIGFSSVLAIVLAVGLTHNKPENVKYGIYLDAGSSHT 60
1 MEDRESELKVCCKNIIILIGFSSVLAIVLAVGLTHNKPENVKYGIYLDAGSSHT 60
1 MEDTRSVKRCCKNIIILIGFSSVLAIVLAVGLTHNKPENVKYGIYLDAGSSHT 60

Db 61 SLIYIKWPAEKENDGVVQLLEECOVKPGISKYAKOTDEIAVLAECMKMTERIPAK 120
61 SLIYIKWPAEKENDGVVQLLEECOVKPGISKYAKOTDEIAVLAECMKMTERIPAK 120
61 SLIYIKWPAEKENDGVVQLLEECOVKPGISKYAKOTDEIAVLAECMKMTERIPAK 120

Db 121 HMEPTVYLGATAGMRLRMESKQADEVLAAYSRSLKSYPFDFOGAKITGOEGAYGWI 180
121 HMEPTVYLGATAGMRLRMESKQADEVLAAYSRSLKSYPFDFOGAKITGOEGAYGWI 180
121 HQEPVYLGATAGMRLRMESKQADEVLAAYSRSLKSYPFDFOGAKITGOEGAYGWI 180

Db 181 TINYLLGKFTQBSWLANIS-DQKQATFGALDGGSSQTQVFPVPLNQTLEAPETSLOFR 239
181 TINYLLGKFTQBSWLANIS-DQKQATFGALDGGSSQTQVFPVPLNQTLEAPETSLOFR 239
181 TINYLLGKFTQBSWLANIS-DQKQATFGALDGGSSQTQVFPVPLNQTLEAPETSLOFR 240

Db 241 LYGDYVYVTSFLCYGQDQALMOKLADIOVASSGILKDPFPGYKVVVNSLQYPR 299
241 LYGDYVYVTSFLCYGQDQALMOKLADIOVASSGILKDPFPGYKVVVNSLQYPR 299
241 LYGDYVYVTSFLCYGQDQALMOKLADIOVASSGILKDPFPGYKVVVNSLQYPR 300

Db 300 CTKREKPLPENOQVOGTGDEYEOCHQSLIKFPNNSHCPSYSCAFNGVFLPPLQSGFAP 359
300 CTKREKPLPENOQVOGTGDEYEOCHQSLIKFPNNSHCPSYSCAFNGVFLPPLQSGFAP 359
301 CTKREKPLPENOQVOGTGDEYEOCHQSLIKFPNNSHCPSYSCAFNGVFLPPLQSGFAP 357

Db 360 SAFYVMDFFPKMANDSVSQEKMETIKNFKSKPWEVKASYPIVYKEKYLSEYCSGTY 419
360 SAFYVMDFFPKMANDSVSQEKMETIKNFKSKPWEVKASYPIVYKEKYLSEYCSGTY 419
361 SAFYVMDFFPKMANDSVSQEKMETIKNFKSKPWEVKASYPIVYKEKYLSEYCSGTY 418

Db 420 ILSLLQGYNFTGSTMDOIHFMWIKIDSNAGWTGLMNLTMIPAEOPLSPPLPSTI 479
420 ILSLLQGYNFTGSTMDOIHFMWIKIDSNAGWTGLMNLTMIPAEOPLSPPLPSTI 479
417 ILSLLQGYNFTGSTMDOIHFMWIKIDSNAGWTGLMNLTMIPAEOPLSPPLPSTI 476

Db 480 YVFLVFLSLVLAAMVTGLTFESKPSYFWKNAV 513
480 YVFLVFLSLVLAAMVTGLTFESKPSYFWKNAV 513
477 YVFLVFLSLVLAAMVTGLTFESKPSYFWKNAV 510

RESULT 3 PRELIMINARY; PRT; 306 AA.
ID 09Y309;
AC 09Y309;
DT 01-NOV-1999 (Tremblrel. 12, Last created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
DE ECTO-ATP DIPHOSPHOHYDROLASE II (EC 3.6.1.5).
GN PLEA 2.

```

Query Match	Best Local Similarity	Score 1931;	DB 4;	Length 306;
Matches 269;	Conservative 3;	Mismatches 6;	Indels 1;	Gaps 1;
<p>OS Homo sapiens (Human). CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; OC Eutheria; Primates; Catarrhini; Hominoidea; Homo. RN [1] RP SEQUENCE FROM N.A. RC TISSUE=PLACENTA; RA MATSUOTO M., SAKURAI Y., KOKUBO T., YAGI H., MATSUI T., TITANI K., RA FUJIMURA Y., NAHITA N.; "The cDNA cloning of human placental ecto-ATP diaphosphohydrolases I and II." RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases. DR EMBL: AJ133134; CAB41887.1; - DR PROSITE: PS01238; GDAL_CD39_NTPASE; 1. KW Hydrolase. SQ SEQUENCE 306 AA; 34175 MW; 2F3C431 CRC32;</p>				
<p>Db 12 KESNVKTFCSNIIIALIGFSSIIIVIALVLAVGLTQNKALPENKYGIVLDAGSSHSLSYI 71 QY 5 KESNVKTFCSNIIIALIGFSSIIIVIALVLAVGLTQNKALPENKYGIVLDAGSSHSLSYI 64 Db 72 YKPAEKENDGVVHOVEBECKVPGGSKFQKQKNEIGYITLDCMEKAREVIPSQHOET 131 QY 65 YKPAEKENDGVVHOVEBECKVPGGSKFQKQKNEIGYITLDCMEKAREVIPSQHOET 124 Db 132 PYVIGATAGMRLMESEELADRVLDIVERSLSNYPPEFGARLITGOEGAYGWITINY 191 QY 125 PYVIGATAGMRLMESEELADRVLDIVERSLSNYPPEFGARLITGOEGAYGWITINY 184 Db 192 LLGKFSQKTRFESIVPEYTNNOETFFGALDGGASTQYTFPQONTISPPNALQFLRYGK 251 QY 185 LLGKFSQKTRFESIVPEYTNNOETFFGALDGGASTQYTFPQONTISPPNALQFLRYGK 244 Db 252 DYNVYTHSLCYGKQDALMOKLADIDIOASTIOS;-RPAPE 289 QY 245 DYNVYTHSLCYGKQDALMOKLADIDIOASTIOS;-RPAPE 283</p>				
<p>RESULT 4 PRELIMINARY; PRT; 493 AA.</p>				
ID 093295;	AC 093295;			
DT 01-NOV-1998 (TREMBlrel. 08, Created)	DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)			
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)	DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)			
DE ECTO-ATP-DIPHOSPHOHYDROLASE.	DE ECTO-ATP-DIPHOSPHOHYDROLASE.			
OS Gallus gallus (Chicken).	OS Gallus gallus (Chicken).			
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.	CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.			
RN [1] RP SEQUENCE FROM N.A. RC TISSUE=OVIDUCT; RX MEDLINE: 98298108. RA NAGY A.K., KNOWLES A.F., NAGAMI G.T.; "Molecular cloning of the chicken oviduct ecto-ATP- diaphosphohydrolase." RL J. Biol. Chem. 273:16043-16049(1998). DR EMBL: AF041355; AAC26491.1; - DR PROSITE: PS01238; GDAL_CD39_NTPASE; 1. DR PFM: PFO1150; GDAL_CD39; 1. KW Hydrolase. SQ SEQUENCE 493 AA; 54034 MW; A6E48E68 CRC32;	RN [1] RP SEQUENCE FROM N.A. RC TISSUE=OVIDUCT; RX MEDLINE: 98298108. RA NAGY A.K., KNOWLES A.F., NAGAMI G.T.; "Molecular cloning of the chicken oviduct ecto-ATP- diaphosphohydrolase." RL J. Biol. Chem. 273:16043-16049(1998). DR EMBL: AF041355; AAC26491.1; - DR PROSITE: PS01238; GDAL_CD39_NTPASE; 1. DR PFM: PFO1150; GDAL_CD39; 1. KW Hydrolase. SQ SEQUENCE 493 AA; 54034 MW; A6E48E68 CRC32;			
<p>Query Match 39.4%; Score 1500; DB 13; Length 493; Best Local Similarity 41.4%; Pred. No. 2,31e-273; Matches 207; Conservative 115; Mismatches 163; Indels 15; Gaps 15;</p>				
<p>Db 5 GKVVAGLITATCVSIIALLISANDVDVLPPTKTKGLVFDAGSTITALYQWPADE 64 QY 14 SKNIIATLIGFSSIIIVIALVLAVGLTQNKALPENKYGIVLDAGSSHSLSYI 72</p>				

D	b	65	NGRTIVSOVSCVNYNGSGISVYADDPAAGASLKPCLDKAMAVIEBOONOTPRYLACTA	124
O	y	73	NDTVSVHQQVEECRCKGKGGISKVFQKVNTEGIYITDCMERAREVIRPSOHQETPYVLGATA	132
D	b	125	GMRLLRFQNSCTKAEQVFAEYSKAIREFPVDFRGAQILITGNEGSPGMITVYVLETLT-K	183
O	y	133	GMRLLRNESELADRYLDYVERLSNTPPDPFOGARITITQDEGATGWITINYLAKPSQK	192
D	b	184	FS-FA-GKWEHPONTELYVALGALDGGASTQITFOP-GVTIEDKNTSVLFRLYGTNYSLYT	240
O	y	193	TRMFSIVPYVE-TNNQETFGALDGGASTQVTFVPQNOTIESPDNALQFRLYGKIDYNYTH	251
D	b	241	SYLCYGTQIOA-SKRIMAAALHODGSYVONISHPCYKPGYRITITAEYDSCVPPPSMLS	299
O	y	252	SFLCYGDKDQALMOKLAKDI-QVASN-EILRDDCFHPGYKKNVNSDLYKTPCTKRFEMTL	309
D	b	300	PAQILVTYTGNGNPAACPITALKLEFNTGCGNRRTCGDDGYOPVPVGGQFAFGFYPSF	359
O	y	310	PPQPFETIQTGNTQOCHQSTLELFPNTSYCPYSOCANFGIFLDPLOGDGAFAFYFWKF	369
D	b	360	LNLTGQO-SLSHVNATVWDFCNKMSSELVEFEPONKE-HLATYCVGLYITLLVDGYKF	417
O	y	370	LNLTSEKVSQEKYTEMKKKRCQAQPMWEIKTSYAGVKEKLYSECSGYITLLDQYHF	429
D	b	418	DEHFWNSIHFSQKAGNADIGWTGFMNLNTNMPIPE-ALEHFKGHEPSLMGAT-SFVYL	475
O	y	430	TADSEWHIHFHIGKQSGDAQWMTGYMLNTNMIPAEQPLSFPLSHSTYVFLMVFSLVF	489
D	b	476	AI-VAGLVAIILLOC-FWKS	493
O	y	490	TVALITGLITFHKPSYFWKDM	509
RESULT		5	PRELIMINARY:	PRT: 495 AA.
ID		055026		
AC		055026:	035928:	
DT		01-JUN-1998	(TREMblrel, 06, Created)	
DT		01-JUN-1998	(TREMblrel, 06, last sequence update)	
DT		01-NOV-1999	(TREMblrel, 12, last annotation update)	
DE			ECO-ATPASE (EC 3.6.1.5).	
GN			CD39JL.	
OS			Mus musculus (Mouse).	
OC			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;	
OC			Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
RN			[1]	
RP			SEQUENCE FROM N.A.	
RC			TISSUE=LIVER.	
RC			MEDLINE: 98289263.	
RA			GAO L., DONG L., WHITLOCK J.P. JR.;	
RT			"A novel response to dioxin. Induction of ecto-ATPase gene	
RT			expression.";	
RL			J Biol. Chem. 273:15358-15365(1998).	
RN			[2]	
RP			SEQUENCE FROM N.A. (SHORT FORM).	
RC			TISSUE=EMBRYO;	
RC			MEDLINE: 97419269.	
RA			CHADWICK B.P., FRISCHAUF A.M.;	
RT			"Cloning and mapping of a human and mouse gene with homology to ecto-	
RT			ATPase genes.";	
RL			Mamm. Genome 8:668-672(1997).	
CC			-1- FUNCTION: HYDROLYSES EXTRACELLULAR ATP TO AMP.	
CC			-1- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 PHOSPHATE.	
CC			-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).	
CC			-1- ALTERNATIVE PRODUCTS: TWO ISOFORMS: A LONG FORM (SHOWN HERE) AND A	
CC			SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.	
CC			-1- INDUCTION: BY DIOXIN.	
CC			-1- PFM: HAS PROBABLY VARIOUS DISULFIDE-BONDS (BY SIMILARITY).	
CC			-1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.	
DR			EMBL: U9542811: AAC24347.1: - - - - -	
DR			EMBL: U91511: AABR1014.1: - - - - -	
DR			MED: MGI:1096863; CD39J1.	
DR			PROSITE; PS01238; GDAL_CD39_NTPASE; 1.	
DR			PFAM: PF01150; GDAL_CD39; 1.	

Query Match	Best Local Similarity	Score 1429:	DB 11:	Length 495:
Matches	214;	Conservative 103;	Mismatches 152;	Indels 21; Gaps 19
KM Hydrolase; Transmembrane; Antigen; Glycoprotein; Alternative splicing.				
FT DOMAIN	1	4	CYTOPLASMIC (POTENTIAL).	
FT TRANSMEM	5	25	POTENTIAL.	
FT DOMAIN	22	25	POLY-LEU.	
FT TRANSMEM	26	462	EXTRACELLULAR (POTENTIAL).	
FT DOMAIN	463	483	POTENTIAL.	
FT TRANSMEM	467	470	POLY-LEU.	
FT DOMAIN	484	495	CYTOPLASMIC (POTENTIAL).	
FT TRANSMEM	130	132	LTS -> MAG (IN SHORT FORM).	
FT VARSPLIC	133	495	MISSING (IN SHORT FORM).	
FT CARBOHYD	64	64	POTENTIAL.	
FT CARBOHYD	129	129	POTENTIAL.	
FT CARBOHYD	294	294	POTENTIAL.	
FT CARBOHYD	319	319	POTENTIAL.	
FT CARBOHYD	378	378	POTENTIAL.	
FT CARBOHYD	443	443	POTENTIAL.	
SO SEQUENCE	495 AA;	54310 MM;	B9CE5702 CRC32;	

[illegible]

RA	KGCEL B.	BRAUN N.,	HEINE P.,	MALISZEWSKI C.R.,	ZIMMERMANN H.;
RT	RNA	ecto-ATPase	and an ecto-ATP	diphosphohydrolase	are expressed in
RT	rat	brain."			
RL	Neuropharmacology	36:1189-1200	(1997).		
CC	- FUNCTION:	HYDROLYSES	EXTRACELLULAR	ATP	TO AMP.
CC	- CATALYTIC ACTIVITY:	ATP + 2 H(2O)	= AMP + 2 PHOSPHATE.		
CC	- SUBCELLULAR LOCATION:	INTEGRAL	MEMBRANE	PROTEIN	(POTENTIAL).
CC	- TISSUE SPECIFICITY:	HIGHEST	LEVELS	IN	HEART, KIDNEY, AND SPLEEN.
CC	STRONG SIGNALS	IN	THYMUS, LUNG,	SKELETAL	MUSCLE, AND BRAIN. VERY
CC	WEAK SIGNAL	IN	LIVER.		
CC	- PTM:	HAS	PROBABLY	VARIOUS	DISULFIDE-BONDS.
CC	- SIMILARITY:	BELONGS	TO	THE	GDAL / CD39 NTPASE FAMILY.
DR	EMBL:	Y11835;	CAA7253.1;		
DR	PROSITE:	PS01238;	GDAL_CD39_NTPASE;	1.	
KW	PFAM:	PF01150;	GDAL_CD39;	1.	
KW	Hydrolase;	Transmembrane;	Antigen;	Glycoprotein.	
FT	DOMAIN	1	4	CYTOPLASMIC (POTENTIAL).	
FT	TRANSMEM	5	25	POTENTIAL.	
FT	DOMAIN	22	25	POLY-LEU.	
FT	DOMAIN	26	462	EXTRACELLULAR (POTENTIAL).	
FT	TRANSMEM	463	483	POTENTIAL.	
FT	DOMAIN	467	470	POLY-LEU.	
FT	DOMAIN	484	495	CYTOPLASMIC (POTENTIAL).	
FT	CARBOHYD	64	64	POTENTIAL.	
FT	CARBOHYD	129	129	POTENTIAL.	
FT	CARBOHYD	284	294	POTENTIAL.	
FT	CARBOHYD	306	306	POTENTIAL.	
FT	CARBOHYD	319	319	POTENTIAL.	
FT	CARBOHYD	378	378	POTENTIAL.	
FT	CARBOHYD	443	443	POTENTIAL.	
SO	SEQUENCE	495	AA;	54389	MM; 24FBF16G CRC32;
Query Match		37.4%;	Score 1422;	DB 11;	Length 495;
Best Local Similarity		44.0%;	Fred. No. 4.05e-257;		
Matches	2133;	Conservative	98;	Mismatches 154;	Indels 19; Gaps 17;

D	b	13	L1AAAGTLL-L-LCVP--TQVREPPALKTGYLDAGSSHTSMFYKKRPADKENDTGYG	70
Q	y	20	ILGFSSIIIVALLAVGLTOKKALPEWKKGIYLDAGSSHTSIYLYKKRPAREKENDGVH	79
D	b	71	OHSSCVOGGGSSVANDPSKAGOSLRCEQLARVY-PRDRASCPPLYLCAFGMPFN	129
Q	y	80	QVEECVAKPGISKVQKVNELTGYLTDCKMRA-REVIIRSOHOETPVYLGATGMKLR	138
D	b	130	L1SPSEATAVLEAVYQTLTQYPDFEFGARILISGODEGVEGWYANYLLENFI-KYGM--V	186
Q	y	139	MESEELADRVLDYVERLSNYPDFQGARILITQEGEAGVGMITIVYLLGKRSQKTRWFSI	198
D	b	187	GRWIRPKCTGLAMDLGASQITFEFTSPS-EDPGNEVHLARYQOHRYVTHSFLCYGR	245
Q	y	199	VPEIETNNQTFEGALDLDGASQTVTFVQNOTIETSPNALQFLYKDKVNYVTHSFLCYGK	258
D	b	246	DDTLRLRLASALQI--HHF-H-PCWPGSYSTOVLQDQVQSCITMGQRPRAFNLSAIVSL	301
Q	y	259	DQALMOKLAKDIQVANSNLLDPCFHHGYKKVNVSDLYKTCYQ-REEMTLPPQQR-EI	316
D	b	302	SGTSNATLCTRDVLSYLFNISSCPSPQSGFNGVQPPVAGNFATFSAFYTYVDFLLTYMGL	361
Q	y	317	QIGIGNYQOCHSIELELFTSTICPIYSQCAFNGITLPLPLQDQFAFAFVFWKFKFLN-LTSE	375
D	b	362	PVGTLKOLEEATETICNQTWTELQARVPGQKTR-LADYCAVMFTHOLLSRGYHDERSF	420
Q	y	376	KYSQKEIVEMMKKF-CQNPWEIKITSYAGVKEKYLEYCEFSSTIYLSLLQGHFTADSM	434
D	b	421	REYVFOKRAADTANGVALGMYLNTNLIPADLEGLKKGTHSSWALLLEFT-VLIATL	479
Q	y	435	EHIIHFGIQSGDAGWTLGYMLNTNNIPAEQF-LSTPLSHSTYVFLMWLFSVLFTVAI	493
D	b	480	VLLL 483	
Q	y	494	IGLL 497	

RESULT	7	PRELIMINARY:	PRT:	494 AA.
ID	P79784			
AC	P79784			
DT	01-MAY-1997 (TREMBLrel. 03, Created)			
DT	01-MAY-1997 (TREMBLrel. 03, Last sequence update)			
DT	01-NOV-1999 (TREMBLrel. 12, Last annotation update)			
DE	ECTO-ATPASE (EC 3.6.1.5) (ATP PYROPHOSPHOROLASE).			
OS	Gallus gallus (chicken).			
OC	Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;			
NC	Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.			
LN	[1]			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RC	TISSUE=SKLETAL MUSCLE, AND GIZZARD.			
RX	MEDLINE; 97150869.			
RA	KIRLEY T.L.;			
RT	"Complementary DNA cloning and sequencing of the chicken muscle ecto-			
RT	ATPase. Homology with the lymphoid cell activation antigen CD39.";			
RL	J. Biol. Chem. 272:1076-1081(1997).			
RN	[2]			
RP	SEQUENCE OF 1-12 AND 154-176.			
RA	STOUT J.G., KIRLEY T.L.;			
RT	"Purification and characterization of the ecto-Mg-ATPase of chicken			
RT	gizzard smooth muscle.";			
RL	J. Biochem. Biophys. Methods 29:61-75(1994).			
CC	-1- FUNCTION: HYDROLYSES EXTRACELLULAR ATP TO AMP.			
CC	-1- CATALYTIC ACTIVITY: ATP + 2 H(2O) = AMP + 2 PHOSPHATE.			
CC	-1- COFACTOR: CALCIUM ION-AND MAGNESIUM ION-DEPENDENT ACTIVITY (BY			
CC	SMILARITY).			
CC	-1- SUBUNIT: HOMODIGOMER.			
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).			
CC	-1- PFM: HAS PROBABLY VARIOUS DISULFIDE-BONDS.			
CC	-1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.			
DR	EMBL; U74467; AAC60071.1; -			
DR	PROSITE; PS01238; GDAL_CD39_NTPASE; 1.			
DR	PFAM; PF01150; GDAL_CD39; 1.			
KM	Hydrolase; Transmembrane; Antigen; Glycoprotein.			
FT	INT_MET 0 0			
FT	DOMAIN 1 3			
FT	TRANSMEM 4 24			
FT	DOMAIN 8 13			
FT	DOMAIN 25 464			
FT	TRANSMEM 465 485			
FT	DOMAIN 486 494			
FT	CARBOHYD 61 61			
FT	CARBOHYD 296 296			
FT	CARBOHYD 417 417			
FT	CARBOHYD 443 443			
FT	CONFLICT 11 12			
FT	CONFLICT 174 176			
SO	SEQUENCE 494 AA; 54402 MW; E4370CCF CRC32;			
	Query Match 37.3%; Score 1418; DB 13; Length 494;			
	Best Local Similarity 40.9%; Pred. No. 2.75e-256;			
	Matches 201; Conservative 122; Mismatches 154; Indels 15; Gaps 14.			
Db	7 VLLLATLGLCLGII-LTCLGSGDARG-PPSEKYGIVDAGSSHTAVTYTKWPAKDENDTG 64			
Oy	17 ILATIGFSSTIAVYALALAVGLTQKALPENKVKYIIVDAGSSHNSLYTYKMPARKENDTG 76			
Db	65 VVSEHSKMDVEGPGISSTYSKPPAGKSLEHCLSQAMRDV-PKEKHADPLTYLGATAGMR 123			
Oy	77 VVHQVEECRVKPGISKRFVKNEIGIYLTDCMERA-REVIPRSQHOETPEYLGATAGMR 135			
Db	124 LTTI-ADPPSQTCLSAVWATLTKSPDPGGKSLTSGEEGVFGTITNNTLYLENIRKG-W 181			
Oy	136 LLRSESELARVLDVVERSLNSNTPFDGKARITITGDEGAYGITTNTLYLGKFSQKTRW 195			
Db	182 LG--EWIISKRTTGLGAMDFGASQTIFET--SDAIEPKNEVMKLGLGOPYKYVTHSFLC 238			
Oy	196 FSIYPIYENNETGALDLDGASQYVTFPQNGNLTIESPDNALQRLTKGXIDYNTVTHSFLC 255			
Db	239 YGRDOVTKRLISKVLAENYQETVIANPCWPTGYRKSLSLSIYDSPCTEKERPGLPINTT 298			

```

QY 256 YKDDALMQLAKLDQVAS -NEILRDPEFHGYYKVVVNSDLYTTPCTCKREMTLPPOQ- 313
Db 299 VVVSQGTGNGNLCAVHNKLFDTSCSFHSCGSDVQPOEVSNGFIASFAYFYVDPIRV 358
QY 314 FEIQIGIYNVOOCHQSLLEIFMTSYCPYSQAFNGFIPLPLOGDGAFAEFYVAKFINLT 373
Db 359 MRPFHPSDLDKAETICATSMNELYOK -APREKLRPDYCATSTFYVLLITGMYFNN 417
QY 374 SKR-V-SOEKVEYEMMKKFCACPWEIKTSYGVAKERYKELYSEYCFGTLYLSLLQGYHFTA 431
Db 418 RSPFIAPCKKAGFSTIGMALGYMLNTFMNMAPDAPBSRMLNMYVILITLIVITLT 477
QY 432 DSWEMIHIFIGKIQSDAGCTIGYMLNTFMNMAPDQPLS -TFLSHSTFYVFLWVLSVLT 490
Db 478 ALLTAVYLLRS 489
QY 491 VAIIGLIFHKP 502

RESULT 8
ID 09Y5L3 PRELIMINARY; PRT; 495 AA.
AC 09Y5L3;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, last annotation update)
DE ECTO-APPASE (EC 3.6.1.3).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominiidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA MATEO J., HARDEN T.K., BOYER J.L.;
RT "Functional expression of a cDNA encoding a human vascular endothelial
RT ecto-APPase."
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL, AF144748; AAD0239.1; -.
KW Hydrolase.
SQ SEQUENCE 495 AA; 53665 MW; 60C0BB61 CRC32;

Query Match 35.5%; Score 1353; DB 4; Length 495;
Match Local Similarity 40.3%; Pred. NC. 8,936-243;
Matches 197; Conservative 117; Mismatches 158; Indels 17; Gaps 16;

Db 6 RSLPLPLLLAAGLAGLLILCVPTEDVRE -PPALKYGIIVLDAGSSHTSMFYKMPADREN 64
QY 15 KILILIDFESS -IAVILLAVGLQNKALPENKYGIVLDAGSSHTSYLYKMPARKEN 73
Db 65 DTGIYGHSSCDYPEGGISSTADNDSGASQSLVGCLEQALODVPERHAGTPLYLGATAG 124
QY 74 DTGVVHQVEECFKVKGPGISKFEVQKNEIGILYITDCMERARFVIRPSHOETPVYLGATAG 133
Db 125 MRLMLTLPBEASTVLMVTHTIQYPPDFDFAARLTSGOEGVGVANTVLENFL -KY 183
QY 134 MRLMSESELADRLVADVERLSWTPDFDFAARLTSGOEGVAGWVITINLLKFEFSOKT 193
Db 184 GWVGR-WERPRK -GTIGAMDIGASTQITFETTSBA -EDRASEVOLHLGYGHVYTHSF 240
QY 194 RMFSTVPEPTNNOEFNALDLDGASTQYTFVPQNGTIESPDNALQFRLYQKDVVYTHSF 253
Db 241 LCYGADQVU -ORLLASA -LQTHGF -H -PCWPRGFSTOVYLGVDVYQSPCTMAQPONFSS 296
QY 254 LCYGADQALMOKLADIDIVASNEILRDPCFHPGKVVVNSDLYTKTCTK --RPEMTLPF 311
Db 297 AAVSLSSGSDPHLCIDVLSGLSPSSCGPESFCPSFGVGVQPPVAAFNVAFAFETVDFLR 356
QY 312 QOPEIQQIGNVOOCHQSLLEIFMTSYCPYSQAFNGFIPLPLOGDGAFAEFYVAKFIN 371
Db 357 TSMGLPVATLQLEAAAVNVCNQTAAQVAPGQORAR -LADYACAGAMFYVOLLRSRYGF 415
QY 372 LTFSE -KVQD -EKVTEEMMKFCAQPMREBELKTSYAGAKVKEKLYSEYCFSGTYILSLQGYTHF 429
Db 416 DERAFGGVIFOKKADTVAVGMAIGYMLNTLIPADPPGLRKGGTDFSSVWVLLLFASAL 475
QY 491 VAIIGLIFHKP 502

```

OY 430 TADSWEHIFIGKIOGSDAGWTIGMLNTMIPAEOP-LSTPLSHSTYVFLMWLFSLV 488
DB 476 LA-ALVLL 483
OY 489 FTVAITIGLL 497

RESULT 9
ID 075355 PRELIMINARY; PRT; 529 AA.
AC 075355;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, last sequence update)
DT 01-MAY-1999 (Tremblrel. 10, last annotation update)
DE CD39L3.
GN CD39L3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97419269.
RA CHADWICK B.P., FRISCHAUER A.M.;
RT "Cloning and mapping of a human and mouse gene with homology to ecto-
ATPase genes.";
RT Mamm. Genome 8:668-672(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96341119.
RA CHADWICK B.P., FRISCHAUER A.M.;
RT "The CD39-like gene family: identification of three new human members
RT the gene family from Drosophila melanogaster.";
RL EMBL; AF039917; AAC39884.1; -
DR PFAM; PF01150; GDAL_CD39; 1.
SQ SEQUENCE 529 AA; 59133 MW; 1AC2F1BA CRC32;

Query Match 35.3%; Score 1344; DB 4; Length 529;
Best Local Similarity 38.9%; Pred. No. 6,60e-241;
Matches 189; Conservative 115; Mismatches 172; Indels 10; Gaps 8;

DB 27 LVVL-LVSIIVLVSTVYIQIHKEVLPRLGKYGIVLDAGSSRTYVYVQMPAEKENNTGV 85
OY 18 LALIGFSSIIAVIALAVGLTKALPENVKYGIIVLDAGSSHTSLYIKMPAEKENDTG 77
DB 86 VSGTFKSVKSGSISSGNPNPDVPRAFECMOKVKGQVPSHLHGSPHILGATAGMRL 145
OY 78 VHOVEECRVKPGISKRVQKVEIGIYLTDCMERAREVIRPSOHOETPVYLGATAGMRL 137
DB 146 RLONETANANVLESISYFESKOPFDFRGAQIISGOEGVYVGTITANYLGNFLEKNIAMH 205
OY 138 RMESELDADVLDVERSLSNYPFDGARIITGOEBGAGWITINILGKFSOKTRMFS 197
DB 206 WV--HPHGVETGALDGGASTQISFVAGEKMDLNTSDIMOVSLYGVYTLTYHSFOCYG 263
OY 198 IVPYETNNQTFGALDGGASTQVTFVPOQNTIESPDNALQFRLYKGDVNVYTHSFLCYG 257
DB 264 RNEAEKFLAMLQNSPTKHLNTPCYPRDYSSFTMGHVFSDCTVDQRPESINPDVI 323
OY 258 KQDALMOKLAKDIOVA-SNEILRDPCHPGYKRVVNSDLYKTPCT--KRFEMLPRQOF 314
DB 324 TFEGTGDPSSCKEKVASIFPKACHOETCSFQGVYOPKIKGPFVAFAGFYTASALNLS 383
OY 315 EIOGIGNYQOCHOSILELFTSYCPYSQ-CAFGNIGFLPRLQDGFAGFSAFYVMKFLNLT 373
DB 384 GS-FSLDTFNSSWTNFCQSNWSQPLLLPKFDEVYARSYCFSANIYHLFVNGYKFTTEET 442
OY 374 SEKVSGEKVTEMMKFCQAOPEMEIKTSYAGVKEKYLSECFSGTYILSLLOGVHTADS 433
DB 443 WPOIHFEKEVGNSSIAMSLGYMLSTNOIPAESPLRLPEPPVFGTIAFTTARLLCT 502
OY 434 WEHIFHIGKIOGSDAGWTIGMLNTMIPAEOP-LSTPLSHSTYVFLMWLFSLV-LFTV 491

DB 503 AFLAYL 508
OY 492 AITIGLL 497

RESULT 10
ID 060495 PRELIMINARY; PRT; 529 AA.
AC 060495;
DT 01-AUG-1998 (Tremblrel. 07, Created)
DT 01-AUG-1998 (Tremblrel. 07, last sequence update)
DT 01-MAY-1999 (Tremblrel. 10, last annotation update)
DE E-TYPE ATPASE.
GN HB6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA SMITH T.M., KIRLEY T.L.;
RL Biochim. Biophys. Acta 0:0-0(1998).
DR EMBL; AF034840; AAC09236.1; -
DR PFAM; PF01150; GDAL_CD39; 1.
SQ SEQUENCE 529 AA; 59190 MW; 14BD39CD CRC32;

Query Match 35.2%; Score 1340; DB 4; Length 529;
Best Local Similarity 38.9%; Pred. No. 4,47e-240;
Matches 189; Conservative 114; Mismatches 173; Indels 10; Gaps 8;

DB 27 LVVL-LVSIIVLVSTVYIQIHKEVLPRLGKYGIVLDAGSSRTYVYVQMPAEKENNTGV 85
OY 18 LALIGFSSIIAVIALAVGLTKALPENVKYGIIVLDAGSSHTSLYIKMPAEKENDTG 77
DB 86 VSGTFKSVKSGSISSGNPNPDVPRAFECMOKVKGQVPSHLHGSPHILGATAGMRL 145
OY 78 VHOVEECRVKPGISKRVQKVEIGIYLTDCMERAREVIRPSOHOETPVYLGATAGMRL 137
DB 146 RLONETANANVLESISYFESKOPFDFRGAQIISGOEGVYVGTITANYLGNFLEKNIAMH 205
OY 138 RMESELDADVLDVERSLSNYPFDGARIITGOEBGAGWITINILGKFSOKTRMFS 197
DB 206 WV--HPHGVETGALDGGASTQISFVAGEKMDLNTSDIMOVSLYGVYTLTYHSFOCYG 263
OY 198 IVPYETNNQTFGALDGGASTQVTFVPOQNTIESPDNALQFRLYKGDVNVYTHSFLCYG 257
DB 264 RNEAEKFLAMLQNSPTKHLNTPCYPRDYSSFTMGHVFSDCTVDQRPESINPDVI 323
OY 258 KQDALMOKLAKDIOVA-SNEILRDPCHPGYKRVVNSDLYKTPCT--KRFEMLPRQOF 314
DB 324 TFEGTGDPSSCKEKVASIFPKACHOETCSFQGVYOPKIKGPFVAFAGFYTASALNLS 383
OY 315 EIOGIGNYQOCHOSILELFTSYCPYSQ-CAFGNIGFLPRLQDGFAGFSAFYVMKFLNLT 373
DB 384 GS-FSLDTFNSSWTNFCQSNWSQPLLLPKFDEVYARSYCFSANIYHLFVNGYKFTTEET 442
OY 374 SEKVSGEKVTEMMKFCQAOPEMEIKTSYAGVKEKYLSECFSGTYILSLLOGVHTADS 433
DB 443 WPOIHFEKEVGNSSIAMSLGYMLSTNOIPAESPLRLPEPPVFGTIAFTTARLLCT 502
OY 434 WEHIFHIGKIOGSDAGWTIGMLNTMIPAEOP-LSTPLSHSTYVFLMWLFSLV-LFTV 491
DB 503 AFLAYL 508
OY 492 AITIGLL 497

RESULT 11
ID 015464 PRELIMINARY; PRT; 472 AA.
AC 015464;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, last sequence update)
DT 01-NOV-1999 (Tremblrel. 12, last annotation update)
DE ECTO-ATPASE (EC 3.6.1.5) (CD39 ANTIGEN-LIKE 1).

GN CD39L1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominoidea; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 97419269.
RA CHADWICK B.P., PRISCHAUF A.-M.;
RT "Cloning and mapping of a human and mouse gene with homology to ecto-
ATPase genes."
RL Mamm. genome 8:668-672(1997).
CC -1- FUNCTION: HYDROLYSES EXTRACELLULAR ATP TO AMP.
CC -1- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 PHOSPHATE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- TISSUE SPECIFICITY: BRAIN, PLACENTA, SKELETAL MUSCLE, KIDNEY,
PANCREAS, HEKT, OVARY, TESTIS, COLON, SMALL INTESTINE, PROSTATE,
AND PANCREAS.
CC -1- PPM: HAS PROBABLY VARIOUS DISULFIDE-BONDS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.
DR EMBL: U91510; AAB81013.1; -.
DR MIM: 602012; -.
DR PROSITE: PS01238; GDAL_CD39_NTPASE; 1.
DR PFM: PFM1150; GDAL_CD39; 1.
KW Hydrolase; Transmembrane; Antigen; Glycoprotein.
FT DOMAIN 1 7
FT TRANSEM 8 28
FT DOMAIN 22 25
FT DOMAIN 29 439
FT TRANSEM 440 460
FT DOMAIN 444 447
FT DOMAIN 461 472
FT CARBOHYD 64 64
FT CARBOHYD 129 129
FT CARBOHYD 294 294
FT CARBOHYD 378 378
FT CARBOHYD 420 420
FT SEQUENCE 472 AA; 51161 MW; 5A924C38 CRC32;
Query Match 29.3%; Score 1115; DB 4; Length 472;
Best Local Similarity 40.5%; Pred. No. 1.58e-193;
Matches 197; Conservative 111; Mismatches 143; Indels 36; Gaps 24;
Db 6 RSLPILLAAAGLAGLILCVPTDRE-PPALKYGVLDASHTSMFLYKPADKN 64
QY 15 KNIILALDFSSI-IAVIALVAGLTQNKALPBNVYGLVLDAGSSHTLYLYKPAEKEN 73
Db 65 DTGIYQHSQCDVPGGIGISSYADNPSGASQSLVGLQALDVPKERRAGTPLYGATAG 124
QY 74 DTGVHQBECRVKPGISKRVQKNEIGYILTCMEBARREVIPIRSQHOEPVYIGATAG 133
Db 125 MRLNLINPEASTSVLMAVHTLTQYPPDFRGARILSGOEGVEGVWTANYLLENFT-KY 183
QY 134 MRLMESEELADRLADVERSLSNYPDFOGARILITGOEGAGYMITINYLKFSQKT 193
Db 184 GWYGR-WRPBRK-GTLGMDLGGASTQITFETSPA-EDRASVQVHLXGQHYRYTHSF 240
QY 194 RMFSTVLPETNNQEFGLADLGASTQVTEVPQNOTLISPDNALQRLYLGADYNYTHSF 253
Db 241 LCYGRDQVL-ORLLASA-LQTHGF-H-PCWPRGSESTVLLDGYOSPCTMAQRPNENSS 296
QY 254 LCYGRDQALQKLMKIDIVASNEILRDCPHRYKVVYVSDIYTPCTK--REMTLPF 311
Db 297 ARVSLSSGSDPHLCRDVLSGLFSSSCPSRCSNGVFPVPAVGNFAFAFYTVDFLR 356
QY 312 QOFELIOGIGNVOOCHOSILFELFNISYCPYSCAFNGIFLPLOGDGFAGAFYFMKFLN 371
Db 357 -TS--MGLP-VA-TL-----Q--OLEAA-A-VN--V---C--NOTMAQQLISRGYFDE 394
QY 372 LTSEKVSQEKYTEMKKCAQDPWEIKTSYAGVEKYISEYCFSGTYLLSLLOGYHFTA 431
Db 395 RAFGVIFQKRAADTVAGMALGYMLNTNLIPADPPGLRKGSTDFSSWVLLLFASALIA 454
QY 432 DSMEIHIFIGIKGSDADAWTIGYMLNTNMLPABQP-LSTPLSHSTYVFLWVLSIVLFT 490

Db 455 -ALVILL 460
QY 491 VAILGEL 497
RESULT 12
ID 09X162 PRELIMINARY; PRT: 483 AA.
AC 09X162;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DE 01-NOV-1999 (Tremblrel. 12, Last annotation update)
DE F7A19.34 PROTEIN.
GN F7A19.34.
OS Arabidopsis thaliana (mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC eudicotyledons; Spermatophytes; Magnoliophyta; eudicotyledons;
OC core eudicots; Rosidae; eustosids II; Brassicales; Brassicaceae;
OC Arabidopsis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV, COLUMBIA;
RA FEEDERSTIEL N.A., PALM C.J., CONWAY A.B., CONN L., HANSEN N.F.,
RA ALFAFI H., ARAUJO R., HUIZAR L., ROWLEY D., BUEHLER E., DUNN P.,
RA GONZALEZ A., KREMENTSKAYA I., KIM C., LENZ C., LI J., LIU S.,
RA LUKOS S., SCHWARTZ J., SHINN P., TORIUMI M., VYSOTSKAIA V.S.,
RA WALKER M., YU G., ECKER J., THEOLOGIS A., DAVIS R.W.,
RL Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AC007576; AAD39311.1; -.
SO SEQUENCE 483 AA; 53425 MW; B134313C CRC32;
Query Match 15.5%; Score 590; DB 10; Length 483;
Best Local Similarity 30.9%; Pred. No. 4.59e-87;
Matches 131; Conservative 104; Mismatches 154; Indels 35; Gaps 25;
Db 67 KLRVSLIDAGSSGTRVAVFGYWFESGKVPDFGKHYANL-KLF-PGLSSYADNPEGAS 124
QY 46 NKVIGIVLDAGSSHTSLIYIK-W-PAEKEN-DTGVHQBECRVKPGISKRVQVQVNEIG 102
Db 125 VSVTKLVEFARQRLPKRFRSRDILMATAGMRL-EVP-VOEQLLETRVLNRSQFM 181
QY 103 IYLPDCMERAREVPIRSQHOEPVYIGATAGMRLMSEELADVLDVVERSLSNYPF 162
Db 182 FRDEMANVSSGSDGISTYITANALGSLG--T--D--PLET-----T-GYELGASAQ 229
QY 163 FQG--ARIITGOEGAGYMITINYLKFSQKTRMFSTVLPETNNQEFGLADLGASTQ 220
Db 230 VTFVSSSEH-V-PPESRTIA-YGNISYTIYSHSPFDYGDKAAL-KLLEKIONSANSTV- 284
QY 221 VTFVPONOTLISPDNALQRLYLGK-DYVNYTHSFCLYGRDQALQKLMKIDIVASNEILR 279
Db 285 DGVEDPCTPGKYIDYDTSKNYSVSGFLADESKLKSLOAENFSKRSRATFALLKEGEN 344
QY 280 DPCFHPGKVVYVSDIYTPCTKRFEMTLPPQOEIOGIGNYQCHOSILEFPTS--Y 337
Db 345 CLYEHCSIGSTFPDLOGSPLATASFYTAFFEL-BEKGWLSLIPAGKRYGGEWSKL 403
QY 338 CPYSCAFNGIFLPLOGDGFAGAFYFMKFLNLTSEKVSQEKYTEMKKCAQDPWEI 397
Db 404 ILEVPTTDEFLRGCFSAATYSMLHDSLGLAID-D-SITYASAGSKHPIPLDALGAF 462
QY 398 KTSYAGVEKYLTSEYCFSGTYLLSLLOGYHFTADSWEHIFIGIKGSD-A-GWTLG-Y 454
Db 463 ILDV 466
QY 455 MLNL 458
RESULT 13
ID 015092 PRELIMINARY; PRT: 609 AA.
AC 015092;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)


```
OY 17 IIAIIIGFSSIIAVIAL-LAVGLTQNKALPENVKYGIULDAGSSHTSLYIKWPAERENDT 75
Db 95 FERRGANYASLKLHPGLSFADDPDQASVSLTELVEFAKGRYPKGMWIEFEVRLMATAGM 154
OY 76 GVVHQVEECRRK-GPGISKFFVQKVNIEIGIYLTDCMERAREVILRSQHQETPVYLGATAGM 134
Db 155 RLLEL-PVQ-EKILGVARRVLKSSGFLPRDEMASVSGSDGVAWVAVANFALGSLG-- 209
OY 135 RLIRMESELADRVLDVVERSLSNYPFDQ--ARITIQEGAGYGMITINYLKFSQK 192
Db 210 G--D--PLKT---T-GIVEIIGASAOVTVS-SEPM-PPEPSRTIS-FGNVTYNLYSH 256
OY 193 TWMFSTIVPYETNNQETFGALDGGASTQVTFYPONOTIESPDNALQFRLYK-DYVNYTH 251
Db 257 SFLHSGQNAH-DKLMGSL-L-SRD--HNSAVEPT-RE--KT---FTDPCARKYNLDAN 305
OY 252 SFLCYGKDQALMOKLAKDIQVASNELRDPCHFPGYKKVYVNSDLYKTPCT-KREMTLP 310
Db 306 YOK-HLSGLA-EESRLS--DSFOAG-GNYSOCRSALFTILODNGRILIIAGFSF-L- 358
OY 311 PQOFELQIGIGNYQOCHQSTILELFNTSYCPYSQCAFNGI-FLPPLQGF-GAFSAFYVMK 368
Db 359 PFGL-GEKAMLSNMTSAGERFCGEDMSKLRKDPSLHEEDLLRYCFSSAYIVSLHDTLG 417
OY 369 FLNLTSEKYSQEKYTEMMKKFCAQPMEEIKTSYAGVKEKYLSEYCFSGTYILSLQGYH 428
Db 418 IPLDD-ERIGYANQAGDIPLDNALGAFIOQTATETSQHNASGNL 460
OY 429 FTADSWEHIFIGKIQSDAGMTLGYMLNTNMIPAEQPLSTPL 472
```

Search completed: Fri May 5 08:58:53 2000
Job time : 504 secs.

QY 121 NYPDFOGARIIITGQEGAGWITINVLGKFSQKTRMFSIVPYETNNQETFGALDLGGA 180
 DB 218 STQVTEVPQNOTIESPDNALQFRLYCKDYNVYTHSFLCYCKDQALMOKLAKDIOVASNEI 277
 QY 181 STQVTEVPQNOTIESPDNALQFRLYCKDYNVYTHSFLCYCKDQALMOKLAKDIOVASNEI 240
 DB 278 LRDPCEHPGKRVKVVNSDLTKPTCKRFEMTLPPQOFEIQIGNYQCHQSILELFTNTSY 337
 QY 241 LRDPCEHPGKRVKVVNSDLTKPTCKRFEMTLPPQOFEIQIGNYQCHQSILELFTNTSY 300
 DB 338 CPYSQCAFNGIFLPLPLOGDFGASAFYFVKFLNLTSEKVSQEKVTEMMKKFCQAPWEEI 397
 QY 301 CPYSQCAFNGIFLPLPLOGDFGASAFYFVKFLNLTSEKVSQEKVTEMMKKFCQAPWEEI 360
 DB 398 KTSYAGVKEKYISEYCSGTYIISLLQGYHFTADSWEHIFHGKIQGSDAGWTLGYMLN 457
 QY 361 KTSYAGVKEKYISEYCSGTYIISLLQGYHFTADSWEHIFHGKIQGSDAGWTLGYMLN 420
 DB 458 LTNNIPAEQPLSTPLSHST 476
 QY 421 LTNNIPAEQPLSTPLSHST 439

RESULT 2

ID W04264 standard; Protein: 510 AA.
 AC W04264;
 DT 29-MAY-1997 (first entry)
 DE Human CD39 protein.
 KW Human; lymphocyte activation marker; gene therapy;
 KW ATP diphosphohydrolase; transgenic animal; pig; porcine; graft;
 KW donor; transplantation; endothelial cell; prosthetic device;
 KW platelet aggregation; inhibition; intravascular.
 OS Homo sapiens.
 PN W09630532-A1.
 PD 03-OCT-1996.
 PF 22-MAR-1996; E01270.
 PR 24-MAR-1995; US-410371.
 PR 12-FEB-1996; US-600383.
 PA (NEMO-) NEW ENGLAND DEACONESS HOSPITAL.
 PA (SANO) SANDOZ LTD.
 PI Bach FH, Rodson S;
 DR WPI: 96-455377/45.
 DR N-PEDB: T33966.
 PT Gene therapy of inflammatory or immunological stimulation of
 PT platelet aggregation - using CD39 protein with ATP
 PT di:phospho:hydrolase activity, useful for preventing or alleviating
 PT thrombotic condition in mammalian subject
 PS Claim 3; Page 39; 65pp; English.
 CC Non-human transgenic or somatic recombinant mammals, whose cells
 CC contain a heterologous DNA encoding a polypeptide (especially human
 CC CD39 protein) having ATP-diphosphohydrolase activity under cellular
 CC activating conditions is claimed. In particular the animal is a pig
 CC and its cells (or tissues or organs) can be used for transplantation.
 CC DNA coding for human CD39 is also useful for genetically modifying a
 CC mammalian cell to render it less susceptible to an inflammatory or
 CC immunological stimulus and platelet aggregation. The modified cells
 CC can be used to prevent or alleviate a thrombotic condition.
 CC The present sequence is that of the human CD39 protein and was
 CC disclosed in J. Immunol. 153 (8) (1994) 3574-3584.
 SQ Sequence 510 AA;

Query Match 100.0%; Score 3275; DB 1; Length 510;
 Best Local Similarity 100.0%; Pred. No. 0.00e+00;
 Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 38 TONKALPENVKYGIYDAGSSHTSLYIKWPAKENDTGVAHVECRVKGPGISKRVOK 97
 QY 1 TONKALPENVKYGIYDAGSSHTSLYIKWPAKENDTGVAHVECRVKGPGISKRVOK 60
 DB 98 VNEIGIYTLDCERAREVIPRSQHOTPYLGTAGMRLIMSEELADRYLDVVERSL 157
 QY 61 VNEIGIYTLDCERAREVIPRSQHOTPYLGTAGMRLIMSEELADRYLDVVERSL 120

DB 158 NYPDFOGARIIITGQEGAGWITINVLGKFSQKTRMFSIVPYETNNQETFGALDLGGA 217
 QY 121 NYPDFOGARIIITGQEGAGWITINVLGKFSQKTRMFSIVPYETNNQETFGALDLGGA 180
 DB 218 STQVTEVPQNOTIESPDNALQFRLYCKDYNVYTHSFLCYCKDQALMOKLAKDIOVASNEI 277
 QY 181 STQVTEVPQNOTIESPDNALQFRLYCKDYNVYTHSFLCYCKDQALMOKLAKDIOVASNEI 240
 DB 278 LRDPCEHPGKRVKVVNSDLTKPTCKRFEMTLPPQOFEIQIGNYQCHQSILELFTNTSY 337
 QY 241 LRDPCEHPGKRVKVVNSDLTKPTCKRFEMTLPPQOFEIQIGNYQCHQSILELFTNTSY 300
 DB 338 CPYSQCAFNGIFLPLPLOGDFGASAFYFVKFLNLTSEKVSQEKVTEMMKKFCQAPWEEI 397
 QY 301 CPYSQCAFNGIFLPLPLOGDFGASAFYFVKFLNLTSEKVSQEKVTEMMKKFCQAPWEEI 360
 DB 398 KTSYAGVKEKYISEYCSGTYIISLLQGYHFTADSWEHIFHGKIQGSDAGWTLGYMLN 457
 QY 361 KTSYAGVKEKYISEYCSGTYIISLLQGYHFTADSWEHIFHGKIQGSDAGWTLGYMLN 420
 DB 458 LTNNIPAEQPLSTPLSHST 476
 QY 421 LTNNIPAEQPLSTPLSHST 439

RESULT 3

ID W04339 standard; Protein: 20 AA.
 AC W04339;
 DT 29-DEC-1996 (first entry)
 DE ATP diphosphohydrolase fragment from pig pancreas.
 KW ATP diphosphohydrolase; ATPase; bovine aorta; pig pancreas;
 KW apyrase; CD39; lymphoid cell activation antigen; enzyme;
 KW platelet aggregation; thrombogenicity; anti-haemostatic.
 OS Sus scrofa.
 PN W09632471-A2.
 PD 17-OCT-1996.
 PF 10-APR-1996; CA0223.
 PR 10-APR-1995; US-419204.
 PA (UYSH) UNIV SHERBROOKE.
 PI Beaudoin AR, Sevigny J;
 DR WPI: 96-477122/47.
 DR isolated ATP di:phospho:hydrolase enzymes - have anti-haemostatic
 PT activity, useful for reducing platelet aggregation and
 PT thrombogenicity
 PS Claim 3; Page 46; 60pp; English.
 CC The bovine aorta (W04335 to W04338 and W04340) and porcine
 CC pancreatic (W04339) ATPases have been partially sequenced.
 CC The sequences have been found to be highly homologous to a human
 CC lymphoid cell activation antigen designated CD39 (Maliszewski et
 CC al. (1994). J. Immunol.: 3574-3583). The complete sequences of the
 CC ATPases types I and II have not been obtained yet. Assuming that
 CC the CD39 gene product is an ATPase type II, the use of CD39
 CC in the reduction of platelet aggregation and of thrombogenicity
 CC may be contemplated, as well as a process of making ATPases
 CC using the CD39 sequence (T38516).
 SQ Sequence 20 AA;

Query Match 3.2%; Score 104; DB 1; Length 20;
 Best Local Similarity 78.9%; Pred. No. 1.03e+01;
 Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

DB 2 SPTQETFGALDLGASTOV 20
 QY 166 TNNQETFGALDLGASTOV 184

RESULT 4

ID W20122 standard; Protein: 248 AA.
 AC W20122;
 DT 07-JUL-1997 (first entry)
 DE H. pylori surface membrane protein, 12969218.aa.
 KW Cytoplasmic; vaccine; prevention; infection; identification;
 KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;

FT	misc_difference	269	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	285	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	312	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	319	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	344	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	358	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	378	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	386	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	465	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	473	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	473	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	494	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	555	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	593	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	602	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	609	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	621	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	724	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	736	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	739	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	786	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	841	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	924	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	934	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	1017	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	1054	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	1127	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	1147	/note-	"corresponds to stop codon in DNA sequence"
FT	misc_difference	1178	/note-	"corresponds to stop codon in DNA sequence"
PN	W09632476-A1.			
PD	17-OCT-1996.			
PF	12-APR-1996; CA0214.			
PR	13-APR-1995; US-421701.			
PA	(MOUN) MOUNT SINAI HOSPITAL CORP.			
PI	Labes M, Lozano A, Roach A, Roder J;			
DR	WPI; 96-477127/47.			
N-PSDB:	T38484.			
PT	Assay for substance that modulates response of neuronal cells - and/or neurite growth associated protein, Petrin, useful in conditions involving nerve damage resulting from traumatic injury, stroke or CNS degenerative disorders			
PS	Claim 9, Page 57-61; 119P; English.			
CC	Rat petrin (W04326) is a protein involved in modulating neurite growth inhibition. The amino sequence was deduced from a cDNA			

CC clone (T38484) derived from an adult rat brain cDNA library; no
 CC coding sequence was indicated. Petrin is a new member of the
 CC protein phosphatase 2C family, and is expressed in neurons in brain
 CC tissue, partic. in the Purkinje cells of the cerebellum. Petrin,
 CC and antibodies raised against it, can be used to modulate neurite
 CC growth and axonal regeneration.
 SQ Sequence 1196 AA;

Query Match 3.2%; Score 106; DB 1; Length 1196;
 Best Local Similarity 31.4%; Pred. No. 7,33e+00;
 Matches 16; Conservative 15; Mismatches 18; Indels 2; Gaps 2;

Db 503 KYEGSHCVQLFQGSFLPISCPLEFISFCGGGVRSFQGFVFL-FL 552
 OY 284 NQQCHOSTLEFNTSYCPYSCAFN-GIFLPLQDGFCAFSAFYVMKFL 333

RESULT 6
 ID W20877; standard; Protein; 95 AA.

DT 18-JUL-1997 (first entry)
 DE H. pylori cytoplasmic protein, 136e10215orf82.
 KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
 KW identification; binding compound; bacterium; life cycle; activator;
 KW bacteria; inhibitor; duodenal ulcer disease; chronic gastritis;
 KW diagnosis.
 OS Helicobacter pylori.
 PN M09640893-A1.
 PD 19-DEC-1996.
 PF 06-JUN-1996; U09122.
 PR 07-JUN-1995; US-487032.
 PR 01-APR-1996; US-630405.
 PA (ASTR) ASTRA AB.
 PI Berglindh OT, Smith D, Mellgaerd BL;
 DR WPI: 97-052306/05.
 DR N-PSDB: T68130.

PT Helicobacter pylori nucleic acid sequences and related
 PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
 PT infection, and to detect Helicobacter
 PS Claim 61; Page 1278-79; 1481pp; English.
 CC This sequence represents a H. pylori cytoplasmic protein.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds.
 CC useful as potential H. pylori life cycle activators or inhibitors.
 CC The genomic sequence of H. pylori (ATCC 55679) was determined from
 CC overlapping contigs generated by mechanically shearing the bacterial
 CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,
 CC and the predicted coding regions defined by computer evaluation. To
 CC identify likely H. pylori antigens for vaccine development, the amino
 CC acid sequences predicted from various ORF were analysed for significant
 CC homology to other known or exported membrane proteins. Having identified
 CC and determined the sequences of interest, particular regions can be
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide
 CC production, e.g. in E. coli hosts.
 SQ Sequence 95 AA;

Query Match 3.1%; Score 103; DB 1; Length 95;
 Best Local Similarity 34.0%; Pred. No. 1,21e+01;
 Matches 18; Conservative 16; Mismatches 14; Indels 5; Gaps 5;

Db 24 GYM-VSLGALKRPTCTNRFYKALLFAIF-YHAVNNFLTQCPHGVREFFSS 74
 OY 249 GYKKVNVSDLYKPTCKRFEM-TLPFOQFEIQTGNTY-QQCH-QSILELFWNT 298

RESULT 7
 ID R85880; standard; Protein; 439 AA.

DT 13-SEP-1996 (first entry)
 DE WD-40 domain-contg. YCUT protein.
 KW WD40 repeat region; beta-transducin; protein-protein interaction; drug;
 KW intracellular signalling; protein kinase C; homology; motif; modulator;
 KW receptors of activated protein kinase; enzyme activity; isozyme; human.

OS Synthetic.
 PN W09521252-A2.
 PD 10-AUG-1995.
 PF 31-JAN-1995; U01210.
 PR 01-FEB-1994; US-190802.
 PA (STRD) UNIT LELAND STANFORD JUNIOR.
 PI Mochly-Rosen D, Ron D;
 DR WPI: 95-283772/37.
 PT New WD-40 (beta-transducin)-derived polypeptide(s) - which alter the
 PT activity of a protein, eg. protein kinase C, which interacts with a
 PT protein contg. a WD-40 region.
 PS Example 5; Page 166-168; 351pp; English.
 CC Proteins R85851-92 are protein in which contain at least one WD-40 (also
 CC called beta-transducin homologous) amino acid repeat motifs. The WD-40
 CC regions are involved in protein-protein interactions between proteins
 CC involved in intracellular signalling. An example of such an interaction
 CC is between protein kinase C and receptors of activated protein kinase
 CC (RACK), esp. RACK-1 (R85850). Proteins R85851-82 were isolated based on
 CC homology with beta-transducin, whereas proteins R85882-92 were isolated
 CC based on homology with the WD-40 consensus sequence (R85893). The
 CC proteins were used to construct the peptides R84928-R85063 and
 CC R85786-R85842. The peptides can be used to identify target proteins
 CC contg. WD-40 motifs, as modulators of enzyme esp. isozyme, activity of
 CC proteins involved in protein-protein interaction and to screen for drugs
 CC that will affect protein-protein interaction involving WD-40 domains.
 SQ Sequence 439 AA;

Query Match 3.1%; Score 102; DB 1; Length 439;
 Best Local Similarity 26.7%; Pred. No. 1,43e+01;
 Matches 16; Conservative 16; Mismatches 23; Indels 3; Gaps 3;

Db 180 AYSPPDSRVYVASEDKIKVMDITSGFCLATEEHTSSVTAVQFARKQVEMSS-SLDGT 238
 OY 123 PFDQGARITITGEGAY-GW-ITIVYLKRSQKTRMWSIVPYETNNQDFGALDGA 180

RESULT 8
 ID W20482; standard; Protein; 70 AA.

DT 15-JUL-1997 (first entry)
 DE H. pylori cytoplasmic protein 4035262.aa.
 KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;
 KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
 KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope.
 OS Helicobacter pylori.
 PN M09640893-A1.

PD 19-DEC-1996.
 PF 06-JUN-1996; U09122.
 PR 07-JUN-1995; US-487032.
 PR 01-APR-1996; US-630405.
 PA (ASTR) ASTRA AB.
 PI Berglindh OT, Smith D, Mellgaerd BL;
 DR WPI: 97-052306/05.
 DR N-PSDB: T67645.

PT Helicobacter pylori nucleic acid sequences and related
 PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
 PT infection, and to detect Helicobacter
 PS Claim 61; Page 648-649; 1481pp; English.
 CC The present sequence shows a Helicobacter pylori cytoplasmic protein
 CC that may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds,
 CC useful as potential H. pylori life cycle activators or inhibitors.
 CC The genomic sequence of H. pylori (ATCC 55679) was determined from
 CC overlapping contigs generated by mechanically shearing the bacterial
 CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,
 CC and the predicted coding regions defined by computer evaluation. To
 CC identify likely H. pylori antigens for vaccine development, the amino
 CC acid sequences predicted from various ORF were analysed for significant
 CC homology to other known or exported membrane proteins. Having identified
 CC and determined the sequences of interest, particular regions can be
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide
 CC production, e.g. in E. coli hosts.
 SQ Sequence 70 AA;

PA (UICL) UNIV ST LOUIS.
PI Chinadurai G;
DR WPI; 99-152099/13.
PM Polymers that bind to anti-oncotic proteins: a useful for

DR WPI; 99-152099/13. ,
PT Polypeptides that bind to anti-apoptotic proteins - useful for

PR protecting against cell death induced by viral infection and to
 PR modulate response to physical and chemical stimuli
 PS Example 6: Column 39-42; 41pp: English.
 CC The present invention describes: (1) a method for regulating cell death,
 CC comprising exposing an isolated cell to a polypeptide selected from
 CC Nip1, Nip2, Nip3, Bip1a, Bip5 and Bip13; (2) a method for neutralising
 CC the activity of the adenovirus E1B 19 kD protein, the Bcl-2 protein or
 CC the BHRF-1 protein, comprising exposing an isolated cell to a
 CC polypeptide as in (1); and (3) a method for detecting molecules that
 CC bind to at least one polypeptide as in (1), comprising lysing cells,
 CC exposing the lysate to the polypeptide and detecting any molecule-
 CC polypeptide aggregates. The methods are useful for providing proteins
 CC able to bind to other proteins known to regulate cell survival e.g. it
 CC is known that E1B 19K protein provides a survival function similar to
 CC the cellular protooncogene bcl-2 gene product which is able to block
 CC apoptosis in haematopoietic B and T cells. The present sequence
 CC represents an adenovirus mutant 19K protein from the present invention.
 SQ Sequence 147 AA:

Query Match 2.8%; Score 93; DB 1; Length 147;
 Best Local Similarity 29.1%; Pred. No. 6.22e+01;
 Matches 16; Conservative 13; Mismatches 24; Indels 2; Gaps 2;
 Db 65 INLGHQALFOEKVKTLD-FSTPGRAAAVAFLSFIDKMSSETHLSGGYLDLDFL 118
 ||| : : |||| : | :
 QY 333 LNLTSSEKYSQEKVTMMKMKFCQAPMEIKTSYAG-VKERYLSFYCFSGTYITSL 386

Search completed: Fri May 5 09:18:10 2000
 Job time : 157 secs.

M P E R C H
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPerch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Fri May 5 09:35:11 2000; MasPar time 13.74 Seconds

Tabular output not generated. 414.072 Million cell updates/sec

Title: >US-09-374-586-2
Description: (1-439) from US09374586.pep
Perfect Score: 3275
Sequence: 1 TONKALPENVKYGVLDAGS.....NITNMIPAEQPLSTPLSHST 439

Scoring table: PAM 150
Gap 11

Searched: 131253 segs, 12956647 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-Issued
1:5A_COMB 2:5B_COMB 3:PC9_COMB 4:backfiles1

Statistics: Mean 33.222; Variance 147.681; scale 0.225

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	102	3.1	439	1 US-08-190-	Sequence 65, Applicati	7.00e+00
2	96	2.9	319	1 US-08-413-	Sequence 119, Applicat	1.86e+01
3	96	2.9	459	1 US-08-220-	Sequence 12, Applicati	1.86e+01
4	96	2.9	459	1 US-08-220-	Sequence 14, Applicati	1.86e+01
5	96	2.9	459	1 US-08-413-	Sequence 12, Applicati	1.86e+01
6	96	2.9	459	1 US-08-413-	Sequence 14, Applicati	1.86e+01
7	95	2.9	1285	1 US-07-582-	Sequence 2, Applicatio	2.19e+01
8	95	2.9	1285	1 US-08-453-	Sequence 2, Applicatio	2.19e+01
9	93	2.8	147	2 US-08-408-	Sequence 15, Applicati	3.02e+01
10	93	2.8	175	2 US-08-408-	Sequence 14, Applicati	3.02e+01
11	93	2.8	175	2 US-08-408-	Sequence 11, Applicati	3.02e+01
12	93	2.8	175	2 US-08-408-	Sequence 10, Applicati	3.02e+01
13	93	2.8	175	2 US-08-408-	Sequence 9, Applicatio	3.02e+01
14	93	2.8	175	2 US-08-408-	Sequence 8, Applicatio	3.02e+01
15	92	2.8	1025	2 US-08-530-	Sequence 23, Applicati	4.85e+01
16	92	2.7	205	4 5175383-6	Patent No., 5175383.	6.64e+01
17	88	2.7	337	2 US-08-806-	Sequence 2, Applicatio	5.68e+01
18	88	2.7	352	2 US-08-751-	Sequence 36, Applicati	6.64e+01
19	88	2.7	361	2 US-08-415-	Sequence 31, Applicati	7.75e+01
20	87	2.7	440	2 US-08-485-	Sequence 27, Applicati	7.75e+01
21	87	2.7	440	2 US-08-485-	Sequence 4, Applicatio	7.75e+01
22	87	2.7	474	1 US-08-459-	Sequence 5, Applicatio	7.75e+01
23	87	2.7	678	1 US-08-288-	Sequence 5, Applicatio	7.75e+01

RESULT	ID	US-08-190-802A-65	STANDARD;	PRT;	439 AA.
24	87	2.7	1089	1	US-08-180- Sequence 36, Applicati 7.75e+01
25	87	2.7	1089	2	US-08-475- Sequence 4, Applicati 7.75e+01
26	87	2.7	1089	1	US-08-460- Sequence 4, Applicati 7.75e+01
27	87	2.7	1089	3	PCT-US92-0 Sequence 4, Applicatio 7.75e+01
28	87	2.7	1089	2	US-08-460- Sequence 36, Applicati 7.75e+01
29	87	2.7	1089	1	US-08-477- Sequence 16, Applicati 5.68e+01
30	87	2.7	1089	2	US-08-487- Sequence 7, Applicatio 1.06e+02
31	87	2.7	1089	1	US-07-991- Sequence 7, Applicatio 1.06e+02
32	85	2.6	163	1	US-08-107- Sequence 2, Applicatio 1.06e+02
33	85	2.6	163	3	PCT-US94-0 Sequence 1, Applicatio 9.05e+01
34	85	2.6	263	1	US-08-351- Sequence 6, Applicatio 1.06e+02
35	85	2.6	417	1	US-08-351- Sequence 7, Applicatio 1.06e+02
36	85	2.6	417	1	US-07-939- Sequence 10, Applicati 9.05e+01
37	85	2.6	423	1	US-07-939- Sequence 12, Applicati 1.06e+02
38	85	2.6	617	4	US-08-137- Sequence 25, Applicati 1.23e+02
39	85	2.6	671	4	US-08-609- Patent No. 5266464. Sequence 13, Applicati 9.05e+01
40	86	2.6	1658	2	US-08-609- Sequence 30, Applicati 9.05e+01
41	86	2.6	1726	2	US-08-609- Sequence 30, Applicati 9.05e+01
42	86	2.6	1726	2	US-08-609- Sequence 30, Applicati 9.05e+01
43	86	2.6	1726	2	US-08-609- Sequence 30, Applicati 9.05e+01
44	86	2.6	1726	2	US-08-609- Sequence 30, Applicati 9.05e+01
45	86	2.6	1726	2	US-08-609- Sequence 30, Applicati 9.05e+01

ALIGNMENTS

Sequence 65, Application US/08190802A
Patent No. 5519003
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: P.O. Box 60850
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-0850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/190,802A
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 8600-0139
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 439 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: Protein
HYPOTHETICAL: NO


```

CC      MOLECULE TYPE: protein
SQ      SEQUENCE 1285 AA; 146565 MW; 8485412 CN;

Query Match      2.9%; Score 95; DB 1; Length 1285;
Best Local Similarity 30.6%; Pred. No. 2.19e+01;
Matches 19; Conservative 13; Mismatches 26; Indels 4; Gaps 4;

D6      823 PYELXPTFFFSGLQVQDGLG-FEBAFAFTRFNFNLVSDRSLMENTMLTLESFDYTPMD 881
OY      302 PYSQCAFNGI-FLPLPGDGFAGFAFYVMKFLN-LTSEKVS-OEKYTEMKKKFCADPWE 358
DB      882 AI 883
OY      359 EI 360

RESULT      8
ID      US-08-453-141-2          STANDARD;          PRT;          1285 AA.
AC      xxxxxx
XX
XX
XX
XX
DE      Sequence 2, Application US/08453141
XX
XX      Sequence 2, Application US/08453141
CC      Patent No. 5885589
CC      GENERAL INFORMATION:
CC      APPLICANT: FOGED, Niels T.
CC      APPLICANT: PETERSEN, Svend
CC      TITLE OF INVENTION: PASTURELLA VACCINE
CC      NUMBER OF SEQUENCES: 2
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: Foley & Lardner
CC      STREET: 3000 K Street, N.W., Suite 500
CC      City: Washington, D.C.
CC      COUNTRY: USA
CC      ZIP: 20007-5109
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/453,141
CC      FILING DATE:
CC      CLASSIFICATION:
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: US 08/293,314
CC      FILING DATE: 22-AUG-1994
CC      APPLICATION NUMBER: US 07/582,945
CC      FILING DATE: 12-OCT-1990
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: PCT/DK89/00084
CC      FILING DATE: 11-APR-1989
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: DK 1995/88
CC      FILING DATE: 04-DEC-1988
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: BENT, Stephen A.
CC      REGISTRATION NUMBER: 29,768
CC      REFERENCE/DOCKET NUMBER: 40369/102/AK20
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (202)672-5300
CC      TELEFAX: (202)672-5399
CC      TELEX: 904136
CC      INFORMATION FOR SEQ ID NO: 2:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 1285 amino acids
CC      TYPE: amino acid
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: protein
SQ      SEQUENCE 1285 AA; 146552 MW; 8486546 CN;

```



```

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/408,095
CC FILING DATE: 21-MAR-1995
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mack, Susan J.
CC REGISTRATION NUMBER: 30,951
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (202)293-7060
CC TELEFAX: (202)293-7060
CC INFORMATION FOR SEQ ID NO: 8:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 175 amino acids
CC TYPE: amino acid
CC STRANDEDNESS:
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 175 AA; 20514 MW; 143104 CN;
SQ
Query Match 2.8%; Score 93; DB 2; Length 175;
Best Local Similarity 29.1%; Pred.No.3.02e+01;
Matches 16; Conservative 13; Mismatches 24; Indels 2; Gaps 2.
Db 65 INLGHQALFOCKVKITLD-FSTPGRAAAVAFLSFIRKDKMSEETHLSGGYLDEL 118
| | : : |||| : | : : : :| | : ||| | : |
QY 333 LNLISEKVSQEKVEMMKFKCAQWEELIKTIYSAG-VKEKIYSEICFSGTIISLT 386
RESULT 12 STANDARD: PRT; 175 AA.
XX US-08-408-095-11
AC xxxxxx
AT
DT
EX
XX
DE Sequence 11, Application US/08408095
XX
CC Sequence 11, Application US/08408095
CC Patent No. 5858678
CC GENERAL INFORMATION:
CC APPLICANT: Chinnadural, Govindaswamy
CC TITLE OF INVENTION: APOPTOSIS-REGULATING PROTEINS
CC NUMBER OF SEQUENCES: 35
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: SUGHRUE, MIOTN, ZINN, MACPEAK & SEAS
CC STREET: 2100 Pennsylvania Avenue, N.W.
CC CITY: Washington
CC STATE: D.C.
CC COUNTRY: USA
CC ZIP: 20037
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/408,095
CC FILING DATE: 21-MAR-1995
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mack, Susan J.
CC REGISTRATION NUMBER: 30,951
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (202)293-7060
CC TELEFAX: (202)293-7860
CC INFORMATION FOR SEQ ID NO: 11:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 175 amino acids

```

```

CC      TYPE: amino acid
CC      STRANDEDNESS:
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: protein
CC      SEQUENCE 175 AA; 20396 MW; 142306 CN;
SQ
Query Match          2.8%; Score 93; DB 2; Length 175;
Best Local Similarity 29.1%; Pred. No. 3.02e+01;
Matches 16; Conservative 13; Mismatches 24; Indels 2; Gaps 2
Db      65 INLHGQALFQEKVITLD-FSTPGRAAAAVAFSLFPIKDKMSEETHLSGGYLLDPL 118
      111 : : 1111 : : : : : : : : : : : : : : : : : : : : : : : :
QY      333 INLTSEKVSQEKVETEMMKFCQAPHEIKTYSYAG-VREKYLSEYCFSGTYILSL 386

RESULT 13
ID      US-08-408-095-10      STANDARD;      PRT;      175 AA.
AC      xxxxxx
XX
XX
XX
XX
DE      Sequence 10, Application US/08408095
XX
XX      Patent No. 5858678
CC      GENERAL INFORMATION:
CC      APPLICANT: Chinadural, Govindaswamy
CC      TITLE OF INVENTION: APOPROPSIS-REGULATING PROTEINS
CC      NUMBER OF SEQUENCES: 35
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
CC      STREET: 2100 Pennsylvania Avenue, N.W.
CC      CITY: Washington
CC      STATE: D.C.
CC      COUNTRY: USA
CC      ZIP: 20037
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: PatentIn Release #1.0, Version #1.30
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/408,095
CC      FILING DATE: 21-MAR-1995
CC      CLASSIFICATION: 435
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Mack, Susan J.
CC      REGISTRATION NUMBER: 30,951
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (202)293-7060
CC      TELEFAX: (202)293-7860
CC      INFORMATION FOR SEQ ID NO: 10:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 175 amino acids
CC      TYPE: amino acid
CC      STRANDEDNESS:
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: protein
CC      SEQUENCE 175 AA; 20369 MW; 143135 CN;
SQ
Query Match          2.8%; Score 93; DB 2; Length 175;
Best Local Similarity 29.1%; Pred. No. 3.02e+01;
Matches 16; Conservative 13; Mismatches 24; Indels 2; Gaps 2
Db      65 INLHGQALFQEKVITLD-FSTPGRAAAAVAFSLFIDKMSSEETHLSGGYLLDPL 118
      111 : : 1111 : : : : : : : : : : : : : : : : : : : : : : : :
QY      333 INLTSEKVSQEKVETEMMKFCQAPHEIKTYSYAG-VREKYLSEYCFSGTYILSL 386

RESULT 14
ID      US-08-408-095-9      STANDARD;      PRT;      175 AA.
XX
XX

```

```

AC      xxxxxx
XX
DT
XX      Sequence 9, Application US/08408095
DE
XX      Sequence 9, Application US/08408095
CC      Patent No. 5858678
CC      GENERAL INFORMATION:
CC      APPLICANT: Chinadural, Govindasamy
CC      TITLE OF INVENTION: APOPTOSIS-REGULATING PROTEINS
CC      NUMBER OF SEQUENCES: 35
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
CC      STREET: 2100 Pennsylvania Avenue, N.W.
CC      CITY: Washington
CC      STATE: D.C.
CC      COUNTRY: USA
CC      ZIP: 20037
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: PatentIn Release #1.0, Version #1.30
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/408,095
CC      FILING DATE: 21-MAR-1995
CC      CLASSIFICATION: 435
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Mack, Susan J.
CC      REGISTRATION NUMBER: 30,951
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (202)293-7060
CC      TELEFAX: (202)293-7860
CC      INFORMATION FOR SEQ ID NO: 9:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 175 amino acids
CC      TYPE: amino acid
CC      STRANDEDNESS:
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: protein
CC      SEQUENCE 175 AA; 20402 MW; 143285 CN;
SQ
    Query Match      2.8%; Score 93; DB 2; Length 175;
    Best Local Similarity 29.1%; Pred. No. 3.02e+01;
    Matches 16; Conservative 13; Mismatches 24; Indels 2; Gaps 2;
Db      65 LNLGQALFQEVYITLD-FSTPRGAAVAFLSTIKRKMSSETHLSGGLDFL 118
    |||::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      333 LNLSEKVSQEKVETEMRKFCAPQPEIKTSYAG-VKEKYLSEYCFSGYIISL 366
RESULT 15
ID      US-08-530-792D-23 STANDARD: PRT; 1025 AA.
XX
AC      xxxxxx
XX
DT
XX      Sequence 23, Application US/08530792D
CC      Patent No. 5972680
CC      GENERAL INFORMATION:
CC      APPLICANT: Knowles, W. J.; Guralski, D.; Haigh, W.; Letsinger, J. T.;
CC      APPLICANT: Clairmont, K.; and Hart, J.
CC      TITLE OF INVENTION: Glucose Transporter Vesicle Aminopeptidase
CC      NUMBER OF SEQUENCES: 23
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: Bayer Corporation
CC      STREET: 400 Morgan Lane
CC      CITY: West Haven
CC      STATE: Connecticut

```

```

CC      COUNTRY :  U.S.A.
CC      ZIP:  06516
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE:  3.5" diskette, 1.44 MB Storage
CC      COMPUTER:  Dell Windows 95 PC
CC      OPERATING SYSTEM:  Windows 95
CC      SOFTWARE:  Wordperfect for Windows 6.1
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER:  US/08/530,792D
CC      FILING DATE:  09/19/95
CC      CLASSIFICATION:  435
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER:  08/309,232
CC      FILING DATE:  09/20/94
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME:  Brewer, Alice A.
CC      REGISTRATION NUMBER:  32888
CC      REFERENCE/DOCKET NUMBER:  MW 323P1
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE:  (203) 812-2705
CC      TELEFAX:  (203) 812-5492
CC      INFORMATION FOR SEQ ID NO: 23:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH:  1025 amino acids
CC      TYPE:  amino acid
CC      STRANDEDNESS:  Single
CC      TOPOLOGY:  linear
CC      MOLECULE TYPE:  protein;
CC      ORIGINAL SOURCE:
CC      ORGANISM:  Rattus norvegicus
CC      STRAIN:  Sprague-Dawley
CC      DEVELOPMENTAL STAGE:  adult
CC      TISSUE TYPE:  skeletal muscle
CC      IMMEDIATE SOURCE:
CC      LIBRARY:  Clontech rat skeletal muscle cDNA library in lambda
CC      LIBRARY:  gtl1 and mRNA isolated from rat skeletal muscle
CC      CLONE:  12.1 (from lambda gtl1 library), PCR product clones 5,
CC      CLONE:  334, and KC44.
CC      FEATURE:
CC      NAME/KEY:  complete amino acid sequence for GTVap, long version
CC      IDENTIFICATION METHOD:  translation from cDNA
CC      SEQUENCE  1025 AA; 117200 MW; 5643821 CN;
SQ
Query Match      2.8%;  Score 92;  DB 2;  Length 1025;
Best Local Similarity 25.8%;  Pred. NO. 3,54e+01;
Matches 23;  Conservative 22;  Mismatches 36;  Indels 8;  Gaps 8;
Db      485 INEGATPMEFSEVKEIKELNSYEDFLDARFKTMRKDSLNSHPRISSVOSS-EQIEEM 543
Oy      316 LQGPFGAISAIFYPWK-PLNTLS-EKYSQEKVTEMMK-F-CAQP-WDEIKTSIAGYKER 370
Db      544 FDSLSYFKGASLL-LMLKSY-LSIEDVFOH 570
Oy      371 YLSEYCFSGYTLISLLQGYHFTADSMWH 399

```

 MWSELEH
 (TM)

Release 3 JA John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
 Run on: Fri May 5 09:18:28 2000; Maspar time 44.24 Seconds
 468,081 Million cell updates/sec
 Tabular output not generated.

Title: >US-09-374-586-2
 Description: (1-439) from US09374586.pep
 Perfect Score: 3275
 Sequence: 1 TQNKALPENVKYGVLDAGS.....NLTNMPAEQPLSTPLSHST 439

Scoring table:
 PAM 150
 Gap 11

Searched: 142080 segs, 47172406 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: p1r62
 1:p1r1 2:p1r2 3:p1r3 4:p1r4

Statistics: Mean 49.140; Variance 92.839; scale 0.529

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query %	Match Length	ID	Description	Pred. No.
1	3275	100.0	510	2	I56242 lymphoid cell activat	0.00e+00
2	366	11.2	557	2	T16696 hypothetical protein	1.19e-47
3	360	11.0	630	2	S50463 hypothetical protein	1.89e-46
4	316	9.6	454	2	JC4616 apyrase (EC 3.6.1.5)	9.65e-38
5	280	8.5	455	2	S48859 nucleoside triphospha	9.04e-31
6	265	8.1	518	2	A40732 guanosine-diphospha	6.47e-28
7	266	8.1	1052	2	T04439 hypothetical protein	4.18e-28
8	131	4.0	371	2	T05213 nucleoside-triphospha	1.16e-04
9	127	3.9	628	2	A55421 hypothetical protein	4.53e-04
10	125	3.8	597	2	T04954 UDP-glucuronosyltrans	8.86e-04
11	117	3.6	528	2	JN0620 thioedoxin H1 homolo	1.22e-02
12	116	3.5	107	2	G69998 conserved hypothetical	1.68e-02
13	109	3.3	386	2	T12886 conserved hypothetical	1.50e-01
14	108	3.3	669	2	T05212 apyrase (EC 3.6.1.5)	2.03e-01
15	104	3.2	44	2	S63501 conserved hypothetical	6.75e-01
16	104	3.2	296	2	A72239 conserved hypothetical	6.75e-01
17	104	3.2	544	2	T05211 hypothetical protein	6.75e-01
18	104	3.2	1680	2	T01367 hypothetical protein	6.75e-01
19	100	3.1	306	2	G70481 thiamine monophospat	2.17e+00
20	100	3.1	367	2	T00580 tubulin alpha-3 chain	1.21e+00
21	102	3.1	450	2	JN0105 probable histidine--t	1.21e+00
22	102	3.1	494	2	T03774 glucuronosyltransfere	1.21e+00
23	102	3.1	530	2	A48633	

SUMMARIES

RESULT ENTRY	1	ALIGNMENTS
TITLE	I56242	#type complete
ORGANISM	lymphoid cell activation antigen - human	
DATE	02-Jul-1996	#formal_name Homo sapiens #common_name man
DATE	03-Jul-1996	#sequence_revision 02-Jul-1996 #text_change
ACCESSIONS	I56242	
REFERENCE	I56242	
#authors	Maliszewski, C.R.; Delespesse, G.J.; Schoenborn, M.A.; Armitage, R.J.; Fanslow, W.C.; Nakajima, T.; Baker, E.; Sutherland, G.R.; Polidexter, K.; Birks, C.	
#journal	J. Immunol. (1994) 153:3574-3583	
#title	The CD39 lymphoid cell activation antigen. Molecular cloning and structural characterization.	
#cross-references	MUID:95015846	
#accession	I56242	
#status	preliminary; translated from GB/EMBL/DBJ	
##molecule_type	mRNA	
##residues	1-510	#label RES
##cross-references	GB:S73813; NID:9765255; PID:9765256	
SUMMARY	#length 510 #molecular_weight 57964 #checksum 4387	
Query Match	100.0%; Score 3275; DB 2; Length 510;	
Best Local Similarity	100.0%; Pred. No. 0.00e+00;	
Matches	439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Db	38 TQNKALPENVKYGVLDAGSHTSLYIKWPAEKENDGVVHOVECKRGKSGISKYQK 97	transcription initiat 9.06e-01
Qy	1 TQNKALPENVKYGVLDAGSHTSLYIKWPAEKENDGVVHOVECKRGKSGISKYQK 60	transcription initiat 1.62e+00
Db	98 VNEIGIYLPDCMERAREVIPSQHOETPVYLGATAGMKLLMESEELADRVLDVVERSL 157	hypothetical protein 1.21e+00
Qy	61 VNEIGIYLPDCMERAREVIPSQHOETPVYLGATAGMKLLMESEELADRVLDVVERSL 120	guanosine-3',5'-bis(d) 9.06e-01
Db	158 NYPDPGARRITTOEGGAGVWITINVLGKFSQKTRFSTVPEETNNOEFGALDGA 217	probable rela protein 1.21e+00
Qy	121 NYPDPGARRITTOEGGAGVWITINVLGKFSQKTRFSTVPEETNNOEFGALDGA 180	DNA-directed RNA poly 9.06e-01
Db	218 STQYTFVPQONTIESPNNALQFRLYGKDYNYTTSFLCYGQDAQIMOKLADIVASNEI 277	periodic cryptophan p 1.21e+00
Qy	181 STQYTFVPQONTIESPNNALQFRLYGKDYNYTTSFLCYGQDAQIMOKLADIVASNEI 240	transcription activator 2.88e+00
Db	278 LRDPCHPFGYKKVNVSDLYKTPCTKRFEMTLPPQOFEIOGIGVYQOCHOSILEFNTSY 337	transcription regulat 5.08e+00
Qy	241 LRDPCHPFGYKKVNVSDLYKTPCTKRFEMTLPPQOFEIOGIGVYQOCHOSILEFNTSY 300	probable lipid-transf 8.84e+00
Db	338 CPYSCAFNGIFLPPLOGDFCAFSAFYVNMKFLMLTSEKVSQEKVETEMKKFCQAPMEI 397	gtp pyrophosphokinase 8.84e+00

```
QY 301 CPYSCAENGJFLPLQDGFSAFYFWMKPLNLTSEKVSQEKYTEMMKKFCAQPWEI 360
|||||
Db 398 KTSVAGVEKXLEYCEGCTYLLSLLOGYHFTADSMWHIHTKQSGDAGWTGMYLN 457
|||||
QY 361 KTSVAGVEKXLEYCEGCTYLLSLLOGYHFTADSMWHIHTKQSGDAGWTGMYLN 420
|||||
Db 458 LTNMIPAEOPSTPLSHST 476
|||||
QY 421 LTNMIPAEOPSTPLSHST 439
|||||

RESULT 2
ENTRY 2
TITLE 116696 #type complete
ORGANISM hypothetical protein R07E4.4 - Caenorhabditis elegans
DATE #formal_name Caenorhabditis elegans
20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change
20-Sep-1999

ACCESSIONS 116696
REFERENCE 218561
#authors Miller, N.
#submission submitted to the EMBL Data Library, October 1995
#description The sequence of C. elegans cosmid R07E4.
#accession 116696
#status Preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 1-557 ##label MITL
##cross-references EMBL:U03652; NID:g1049390; PID:g1049394;
##experimental_source strain Bristol C52

GENETICS
#gene CESP:R07E4.4
#introns 39/1: 67/1: 145/2: 192/1: 244/3: 317/3: 451/3: 488/2
SUMMARY #length 557 #molecular_weight 63056 #checksum 4267

Query Match 11.2%; Score 366; DB 2; Length 557;
Best Local Similarity 25.6%; Pred. No. 1.19e-47;
Matches 111; Conservative 109; Mismatches 180; Indels 34; Gaps 27;

Db 44 YGVICDAGSTGRFLPVYMWISDSELDQIEPVYIDNKPVMKKISPGTSTKRAQAE 103
|||||
QY 12 YGVICDAGSTGRFLPVYMWISDSELDQIEPVYIDNKPVMKKISPGTSTKRAQAE 66
|||||
Db 104 YLRPLMELAEHRIPEKRPYVFEATAGMRLIPDEVYLQOKBAVLKLNKLPKITS 163
|||||
QY 67 YLTDCMERAREVIRPSQOETPVYLGATAGMLLMESELADR--VLDYVERSLSNY-P 123
|||||
Db 164 MOVLEKHRIIEGKWEGLYSIAVYALGKENKTALDPGTSAPAHAKTVGMTDMGA 223
|||||
QY 124 FD-FQG-ARITGOEGAGYGTITNYLLGKFSQKTRW-FSIVPEYTNQETFGALDLGA 180
|||||
Db 224 SAQIAFELPDPSISINVENINLGRDDSLFKKLFVTTFLLGAVNGIKKYEHLMLLS 283
|||||
QY 181 STQVTF-VPQNOTLES-P-DN-ALQFLRYGK--DYNVYTHSFLCYGKQDALW---OKLAK 231
|||||
Db 284 KLRQNGTVIIDDCCPLMLHKTVTLENG-ENF-VRRGTGNMWTGKNEVYKLN-PESSE 340
|||||
QY 232 DIQVASNILDLDPCHHPKRYKVVNSDLKTPCTRTREFTLRFQOPELQIGINTQOCHS 291
|||||
Db 341 VCRA-EAKCYFAGVAPASIPLSNIEM-YG-FSEYWSTHDVLGGO-YDAENITAKKTQ 396
|||||
QY 292 ILLEFNISYCPYSOCAFNGIFLPLQDGFSAFYFWMKPLNLTSEKVSQEKYTEMMKK 350
|||||
Db 397 QYCKRNSTIOAESKQLYPRADERLRTQCKSKAMITSVLHDGS-VDKTINKQSVST 455
|||||
QY 351 KFCQAPW--E-EIKTS-YAGVKEKYLSEYCSGYTLLSLLOGYHFTADSMWHIHTK 405
|||||
Db 456 IAGQEVQWALGAMI 469
|||||
QY 406 IQGSDAGWTGMYL 419
|||||

RESULT 3
```

```
ENTRY S50463 #type complete
TITLE hypothetical protein YER005w - yeast (Saccharomyces cerevisiae)
ORGANISM #formal_name Saccharomyces cerevisiae
DATE #formal_name Saccharomyces cerevisiae
28-May-1993 #sequence_revision 24-Feb-1995 #text_change
21-Nov-1997

ACCESSIONS S50463
REFERENCE S50428
#authors Dietrich, F.S.
#submission submitted to the EMBL Data Library, December 1994
#description The sequence of S. cerevisiae cosmids 9537, 9581, 9495, 9867,
and lambda clone 5898.
#accession S50463
#molecule_type DNA
#residues 1-630 ##label DIE
##cross-references EMBL:U08778; NID:g603592; PID:g603597; MIPS:YER005w

GENETICS
#map_position 5R
SUMMARY #length 630 #molecular_weight 71851 #checksum 7317

Query Match 11.0%; Score 360; DB 2; Length 630;
Best Local Similarity 28.4%; Pred. No. 1.89e-46;
Matches 106; Conservative 93; Mismatches 137; Indels 37; Gaps 28;

Db 52 IHOEKDWTFKLNPGSLSEPKRQDAYKSHIKPLDFAKNIIPESHWSSCPVFIQATAGMR 111
|||||
QY 41 VHQVECECRVK-GPGISKRVQKVNEL-GIYLDCEMERAREVIRPSQHOETPVYLGATAGMR 98
|||||
Db 112 LIPDIOSSILDGLCGKLKP-AEFLVEDCSAQIOVDIGETEGLYGWLNTLYGHNND- 169
|||||
QY 99 ILRMSESE-LADRVLDYVERSLSNYPF-DFOGA-RITGOEGEGAYGWTITNYLLGKFSOK 155
|||||
Db 170 ---VN--P-EVSDHPTTFPFMOMGASSTOIAFAPDSCGSIANHRDITVIFRSVNGSLQK 223
|||||
QY 156 TRWESIVPETINQETFPALDGGASTQVTVYVPOQT-TESP-DN-ALQF-R-LYGA-D- 208
|||||
Db 224 WDVFSTWLGFGANARRRYLAOLINTLPENT-ND-YENDDEFSTR-NLNDPCMRGSSSTD 280
|||||
QY 209 YVYTHSFLOCYGDQALQKLAQIVASNELRDPCHHPKRYKVVNSDLKYKP-CTKR 267
|||||
Db 281 PEFKDTI--PHIAGSGNTEQCKTSYPLLLKNMPCDDEPCLFNQVNAHRIDFANDKTTGT 338
|||||
QY 268 FEMTLRPFQOFELOGIGNTQOCHOSILFELFNYS-CPYSQCAFNGIFLPLPDQ-G-D-FGAF 323
|||||
Db 339 SEYWTADVETLIGE--NPFKFSKSLREFPCNSMMWTOILANSDDGYNSIPENLTKACF 397
|||||
QY 324 SAFYFVMK-FLNLISEKVSQEKYTEMMKKFCAQPWEI-----KTSVAGVEKXLEYCE 377
|||||
Db 398 KGNVYLNILHDEGF 410
|||||
QY 378 SGTYLLSLLOGY 390
|||||

RESULT 4
ENTRY 4
TITLE Jc4616 #type complete
ALTERNATE_NAMES adenylylpyrophosphatase: ATP-diphosphohydrolase
ORGANISM #formal_name Solanum tuberosum #common_name potato
DATE #formal_name Solanum tuberosum
10-May-1996 #sequence_revision 19-Jul-1996 #text_change
28-May-1999

ACCESSIONS Jc4616
REFERENCE Jc4616
#authors Handa, M.; Guidotti, G.
#journal Biochem. Biophys. Res. Commun. (1996) 218:916-923
#title Purification and cloning of a soluble ATP-diphosphohydrolase
(Apyrase) from potato tubers (Solanum tuberosum).
#cross-references MUID:96158985
#accession Jc4616
#molecule_type mRNA
#residues 1-454 ##label HAN
##cross-references GB:U058597; NID:g1381632; PID:AA02720.1;
PID:g1381633

#accession PC4147
```


	#molecule_type protein	
	#residues 59-95;96-111,132-160;236-253;332-345	#label HA2
	#experimental_source tubers	
	#note The authors translated the codon GCA for residue 215 as Gly	
COMMENT	This enzyme belongs to a family of E-type ATPases, and it catalyzes the hydrolysis of phosphoanhydride bonds of nucleoside tri- and di-phosphates in the presence of divalent cations. This enzyme has nucleotide substrate specificity, divalent cation requirement, and insensitivity to known inhibitors of ATPases, phosphatases, and adenylate kinase. It is a regulator of various steps involved in starch synthesis.	
GENETICS		
CLASSIFICATION	trpO1	
KEYWORDS	#superfamily nucleoside triphosphatase chromatin-associated glycoprotein; hydrolase; transmembrane protein	
FEATURE		
1-30	#domain signal sequence #status predicted #label TMV	
8-25	#domain transmembrane #status predicted #label TMV	
31-454	#product ATP-diphosphohydrolase #status predicted #label MAY	
44-65	#region actin-heat shock protein 70-hexokinase beta-phosphate binding\	
44-65	#region nucleotide binding\	
192-212	#region actin-heat shock protein 70-hexokinase gamma-phosphate binding\	
192-212	#region nucleotide binding #status predicted\	
350-410,427-446	#region hydrophobic carboxyl end\	
151,262	#binding site carbohydrate (Asn) (covalent) #status predicted	
SUMMARY	#length 454 #molecular-weight 50040 #checksum 6727	
Query Match	9.6%; Score 316; DB 2; Length 454;	
Best Local Similarity	29.0%; Pred. No. 9,65e-38;	
Matches	62; Conservative 58; Mismatches 86; Indels 8; Gaps 7;	
Db	42 ESEHYAVIDAGSTGSRVHFRP-DKGLIPLIGNIEFMATEPGLSVAEDPPAAANS 100 : ::::: : + :	
OY	8 ENVRYGVLDSSHTLYIKKPAREKENDGTGVHVECRKVGCSIKFQKVNEIGTY 67 :	
Db	101 LEPIIDGAEVVPQELSETPELGGATAGRMKGDAEKTIQANRNLY-KNQSFFHSKD 159 :	
OY	68 LTDCMERREVRIPISQOEPTPVYLGAAGRLLRMSSE-LADRVLDVERLSNYPDF 126 :	
Db	160 QWVTIIDTQEGSYMAAIAYILGNLG-KD-YKSTT-ATIDLGG-GSVOMAYASNQF 214 :	
OY	127 QGARIITGQEGGACGWTTIVYLGRKSQKRWSIYPYETNNQETFGALDGCASYTF 186 :	
Db	215 AKAPONDEGEPRVVOOKHMSKDYNLVHSLNTYG 248 :	
OY	187 VPQNQTIESPDNALFRLYGKYDVNYTHSFLCYG 220 :	
RESULT	5	
ENTRY	S48859	#type complete
TITLE	nucleoside triphosphatase precursor, chromatin-associated - garden pea	
ORGANISM	formal name Pisum sativum #common.name garden pea 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 23-Sep-1999	
DATE	23-Sep-1999	
ACCSSIONS	S65147; S48859	
REFERENCE	S65141	
#authors	Hsieh, H.L.; Tong, C.G.; Thomas, C.; Roux, S.J. Plant Mol. Biol. (1996) 30:135-147	
#journal	Light-modulated abundance of an mRNA encoding a calmodulin-regulated, chromatin-associated NTPase in pea.	
#title	calmodulin-regulated, chromatin-associated NTPase in pea.	
#cross-references	MOTU:96197404	
#accession	S65147	
#molecule_type	mRNA	
#residues	1-455 #label HS2	
#cross-references	EMBL:D33743; NID:9563611; PTDN:CAA83655.1; PID:9563612	
CLASSIFICATION	#superfamily nucleoside triphosphatase chromatin-associated	

KEYWORDS	nucleus	length 455	#molecular-weight 50072	#checksum 2198
SUMMARY				
Query Match	8.5%;	Score 280;	DB 2;	Length 455;
Best Local Similarity	28.6%;	Pred. No. 9,04e-31;		
Matches	62;	Conservative	57;	Mismatches 85; Indels 13; Gaps 11.
Db	44 YAYVFDAGSTGRHIVHF-NQNIIDLHIGGVEYINKITPGLSSYANMFQAAKSLIPL 102			
Qy	12 YGIYLDAGSHSTSLIYIKWPAEKENDGVVHVGECRCYKGLSKFVQKNEIGIYLTDC 71			
Db	103 LEQAEVYDDLLQKTPRIGATGRLT--NGRA-SKKILOSVDMLSNRSPFNVPDA 159			
Qy	72 MERAREVYPRSOHOETFPYLGATGMRLLRMESELADRVLDVVERSLN-YPPDFQ-GA 129			
Db	160 VSIIIDTQEGSYLWTVYAVYALNGIKKTKVGVYIDLGGSQVQ--AYAVSKTKAK-N-A 215			
Qy	130 -RIITQEGEAGVGMVITINYLKGRSQK-TRRFSLVPEPTNNQELFFGALDLGGAFTQYTFV 187			
Db	216 PKVADGDDP-YIKKVLKGIPIYDLVHSYHLFGREAS 251			
Qy	188 PONGTIESPDNALQFRLYGKDYNYVTHSFLCYGRDOA 224			
RESULT	6			
ENTRY	A40732	#type complete		
TITLE	guanosine-diphosphatase (EC 3.6.1.42) - yeast (<i>Saccharomyces cerevisiae</i>)			
ALTERNATE_NAMES	protein YEL042w			
ORGANISM	#formal_name <i>Saccharomyces cerevisiae</i>			
DATE	21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 06-Feb-1998			
ACCESSIONS	A40732; B40733; S30837; S50502			
REFERENCE	A40732			
#authors	Abeljon, C.; Yanagisawa, K.; Mandon, E.C.; Hausler, A.; Moremen, K.; Hirschberg, C.B.; Robbins, P.W.			
#journal	J. Cell Biol. (1993) 122:307-323			
#title	Guanosine diphosphatase is required for protein and sphingolipid glycosylation in the Golgi lumen of <i>Saccharomyces cerevisiae</i> .			
#cross-references	MUID:93308137			
#accession	A40732			
#molecule_type	DNA			
#residues	1-518 #label ABE			
#cross-references	EMBL:L19560; NID:g349392; PID:g349393			
#note	sequence extracted from NCBI backbone (NCBIN:134708, NCBI:134711)			
#accession	B40732			
#molecule_type	protein			
#residues	125-144;238-257;276-281;366-374;399-412 #label AB2			
REFERENCE	S30812			
#authors	Mulligan, J.T.; Dietrich, F.S.; Hennessey, K.M.; Sehl, P.; Komp, C.; Wei, Y.; Taylor, P.; Nakahara, K.; Roberts, D.; Davis, R.W.			
#submission	submitted to the EMBL Data Library, February 1993			
#accession	S30837			
#molecule_type	DNA			
#residues	1-518 #label MUI			
#cross-references	GB:U18779; EMBL:L10830; NID:g603625; PID:g603637			
REFERENCE	S50491			
#authors	Dietrich, F.S.			
#submission	submitted to the EMBL Data Library, December 1994			
#description	The sequence of <i>S. cerevisiae</i> cosmid 8199, 8334, and 9871.			
#accession	S50502			
#molecule_type	DNA			
#residues	1-518 #label DIE			
#cross-references	EMBL:U18779; NID:g603625; PID:g603637; MIPS:YEL042w			
GENETICS				
#gene	SGD:GDA1			
#cross-references	SGD:S0000768; MIPS:YEL042w			
#map_position	5L			
FUNCTION				
#description	hydrolase			

KEYWORDS glycoprotein; Golgi apparatus; hydrolase; transmembrane protein

FEATURE 10-24 #domain transmembrane #status predicted #label TM\

41,280,335 #binding_site carbohydrate (Asn) (covalent) #status predicted

SUMMARY #length 518 #molecular-weight 56821 #checksum 4379

Query Match 8.1%; Score 265; DB 2; Length 518;

Best Local Similarity 27.8%; Pred. No. 6,47e-28; Mismatches 180; Indels 45; Gaps 37;

Matches 122; Conservative 92; Mismatches 180; Indels 45; Gaps 37;

Db 88 SEEHKVIIDAGSTGSRVHYKFDYCTSPPT-LDEKPD-MLE-PLSSFDSDYGAAN 144
 QY 7 PENVKIGYILDAGSSHTSLYIKWPAKEKNDTCVHQVEECRVKGPISKF-YQKVAEIG 65
 Db 145 S-LDPLIKAMNVPIKANSCTPVAVKATAGLGLDASSKILSAVRHLEKDYR-FPV 202
 QY 66 IYLDICMERAREVIPSQOETFPVYLGATAGMRLT-RMESEELADRYLDVVERSLSNYPF 124
 Db 203 VEGDGVISNGDEGEVFAWTTTYLTGNIGANGP--KL-P-TAA--VE---DLGGGSTQ 252
 QY 125 -DPGARITTGDEBAGVGTITTYLTGKFSQKTRMFSIVPEYETNNQETFGALDGGASTQ 193
 Db 253 IVEEPTFPINEKAVDEHREFDLKFGEENTYLYQSHLYGLKEG-RNKV-NSVLYE-NA- 308
 QY 184 VTFVPOQNTIES-PDNALQFRL-YGKD-YNVYHSFLCGKQDALMQKLAKDIQVANSNEI 240
 Db 309 LKDGLIKGDMTKTHOLSEPCLEPPKVNATINEKYLTSKEYTITDFIGPDESSGAQCFELT 368
 QY 241 LBDP-CFHGKYKKVAVSDLYKTP-CTKRPE-MTLRPOQ-FEIQIGIGNYQC-QOS-IL- 293
 Db 369 DELINKDACQSPSCFNGVHOPSLVTRFKESNDIYFSEYFRTPLGMLSFITNELN 428
 QY 294 -ELENT-STCPYSQCAFNGITFLPLDQGF--GA-FSAF-YFVKFPLNLT-SEKVSQEKVT 346
 Db 429 DLARIYCKGEETWNSVSGISIDELSPSHRCLDLSEFQVSLHTGYDIPLO-RE-LRT 486
 QY 347 EMKKKFC-AQ-PWEIKTYSAGVKEKYLE--YCFSGTYILSLDGYHTADSWEHIF 402
 Db 487 GKRIKKEIGKICLASLPL 505
 QY 403 ICKIGSDAGMTLGYML 421

RESULT 7
 ENTRY T04439 #type complete
 TITLE hypothetical protein T18B16.150 - Arabidopsis thaliana
 ORGANISM #formal_name Arabidopsis thaliana #common_name mouse-ear cress
 DATE 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 30-Apr-1999
 ACCESSIONS T04439
 REFERENCE 215359
 #authors Bevan, M.; Benes, V.; Rechmann, S.; Borkova, D.; Ansoorge, W.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Schueller, C.
 #submission submitted to the Protein Sequence Database, April 1998
 #accession T04439
 #molecule_type DNA
 #residues 1-1052 #label BEV
 #cross-references EMBL:AL021687
 #experimental_source cultivar Columbia; BAC clone T18B16

GENETICS
 #map_position 4
 #introns 79/3; 118/3; 217/1; 295/3; 396/3; 531/2; 815/3
 #note T18B16.150
 SUMMARY #length 1052 #molecular-weight 116388 #checksum 4955

Query Match 8.1%; Score 266; DB 2; Length 1052;
 Best Local Similarity 27.5%; Pred. No. 4.18e-28;
 Matches 104; Conservative 103; Mismatches 125; Indels 46; Gaps 33;

Db 578 PGFDKLVNRTGLTKAIKPLIQAEMKQIPKNAHRTSLFVYATAGVRLR--PAD-SSWI 634

QY 52 PGISKEVQKVNIGIYILTCMERAREVIPSQOETFPVYLGATAGMRLT-RMESEELADRV 111
 Db 635 LGVNSIILAKSPFTCRREWKITISGTEEAFFGWTALNT-----QTSN-LGLP---KK- 683
 QY 112 LDVVERSLSNYPD-FQG-ARIITGQEGAYMTITNYLTGKFSQKTRMFSIVPEYETNNQ 169
 Db 684 ATFGALDGGSLQYTFEERHNETNLTNRIGSVNHLISAVSLAGYLANDAFRSVYH 743
 QY 170 ETFGALDGGASQYTFVFNQNTIESPDNALQPRLYGKRYNTYS-F-L-CYKDDA- 224
 Db 744 LKKLPNVKSDILEKLEMKHPCLNSGVN-GQYICSQG-ASSVQGKKSGSVIKLVG 801
 QY 225 LMOKLAK-DI-QVASNEI-LRQPCFHPGKYKVVNSDLYKTKTRFEMTLRPOQFEIQG 281
 Db 802 APWAGEC--SALAK-NAP-C--AL-P-DG-YPRP-HGQYVANSFGFYVRRFNLSAF-AS 850
 QY 282 IGVYQOCHOSILELFTSYSCYQCAFNGITFLPLDQGFARSAYFVKKFLNLTSEKVS 341
 Db 851 LDVLEKREPCDKAMQVARTS-VSP-QPFIQYCFRAPHYVSLRREGT-YTDDK--QI- 904
 QY 342 QEKVTEMKKFKCAQPEBEIKTSYAGVKEKYLSEYCFSTGYILSLDGYHTADSWEHIF 401
 Db 905 IIGS--GSIT-WTLGYAL 919
 QY 402 FICKIGSDAGMTLGYML 419

RESULT 8
 ENTRY T05213 #type complete
 TITLE hypothetical protein F1715.40 - Arabidopsis thaliana
 ORGANISM #formal_name Arabidopsis thaliana #common_name mouse-ear cress
 DATE 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Apr-1999
 ACCESSIONS T05213
 REFERENCE 215404
 #authors Bevan, M.; Vitale, D.; Laguori, R.; Argiridou, A.; De Simone, V.; Hohenseil, J.; Mewes, H.W.; Mayer, K.F.X.; Schueller, C.
 #submission submitted to the Protein Sequence Database, July 1998
 #accession T05213
 #molecule_type DNA
 #residues 1-371 #label BEV
 #cross-references EMBL:AL031032
 #experimental_source cultivar Columbia; BAC clone F1715

GENETICS
 #map_position 4
 #introns 14/1; 43/3; 117/1; 171/3
 #note F1715.40
 SUMMARY #length 371 #molecular-weight 42304 #checksum 9483

Query Match 4.0%; Score 131; DB 2; Length 371;
 Best Local Similarity 23.8%; Pred. No. 1.16e-04;
 Matches 19; Conservative 29; Mismatches 26; Indels 6; Gaps 6;

Db 197 HOBOSIYKARKGAVRIRAVSEGOPIPNATISIQQRLLG-PFPGCEVEKNILG-NQ-AYE 253
 QY 84 HGETPYTLCATAGMKRLIRNESE-E-LADRYLDVVERSLSNYPDQGARITTGDEBAGY- 140
 Db 254 NMFOTRFVYTFPANEKMYTS 273
 QY 141 GMTITNYLTGKFSQKTRMFS 160

RESULT 9
 ENTRY A55421 #type complete
 TITLE nucleoside-triphosphatase (EC 3.6.1.15) 1 precursor -
 ORGANISM Toxoplasma gondii
 DATE #formal_name Toxoplasma gondii
 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 17-Mar-1999
 ACCESSIONS A55421
 REFERENCE A55421

```
#authors      Bernhardt, D.; Peck, K.R.; Affifi Affifi, M.; Beckers, C.J.M.,  
               Joiner, K.A.  
#journal      J. Biol. Chem. (1994) 269:29252-29260  
#title        Randomly repeated genes encode nucleoside triphosphate  
               hydrolyase isoforms secreted into the parasitophorous  
               vacuole of Toxoplasma gondii.  
#cross-references MIMD:95050750  
#accession    A55421  
##status      preliminary  
##molecule_type DNA  
##residues     1-628 ##label BEP  
##cross-references GB:U14322; NID:9562078; PID:9562079  
KEYWORDS      hydrolase  
SUMMARY       #length 628 #molecular-weight 69585 #checksum 5865
```

```
Query Match          3.9%; Score 127; DB 2; Length 628;  
Best Local Similarity 32.8%; Pred.No. 4,53e-04;  
Matches            22; Conservative 21; Mismatches 22; Indels 2; Gaps 2;
```

Db 228 TRPTGAEFGFPAFTLTNHSRLRGEDPARCMIDEYGVKHCNRDLAAGEVEYGASAOIVF 287
 | | | | | : | : | : | : : : : | : : : : | : : : : | : : : : |
Qy 129 ARITGGEGAGVGTWITNYLLGFKFSOKTRMPISVPETNN-QETF--CALDIGASTGYTF 186

Db 288 PLEQCTV 294
 | : | : |
Qy 187 VPQNQT I 193

```
RESULT         10  
ENTRY           T04954             #type complete  
TITLE           hypothetical protein F7J7.170 - Arabidopsis thaliana  
ORGANISM        #formal_name Arabidopsis thaliana #common_name mouse-ear  
                cress  
DATE            23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change  
                11-Jun-1999  
ACCESSIONS      T04954  
REFERENCE        Z15391  
AUTHORS         Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.;  
                Mewes, H.W.; Meyer, K.F.X.; Scheller, C.  
#submission     submitted to the Protein Sequence Database, July 1998  
#accession      T04954  
##molecule_type DNA  
##residues      1-597 ##label BEV  
##cross-references EMBL:AIO21960  
##experimental_source cultivar Columbia; BAC clone F7J7
```

```
GENETICS  
#map_position   4  
#introns        F7J7.170  
#note           #length 597 #molecular_weight 66904 #checksum 5095
```

SUMMARY

```
Query Match          3.8%; Score 125; DB 2; Length 597;  
Best Local Similarity 26.8%; Pred.No. 8,86e-04;  
Matches            26; Conservative 28; Mismatches 38; Indels 5; Gaps 5;
```

Db 412 RDLAKSSNVLDLOMLPKRISDFGMAR-OFDENPGOAVTVRRVGYGYAPAMHGRSVK 470
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 98 RLRLMESELADRYL-DVERVSLSNYPDPGALGITIGQEBSGAATWITNYLL-GKSÖK 155

Db 471 TDVYSFGLVLIEITGKRNSGLIGEGETDLPTFAWM 507
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 156 TR-W-FSIVPYETNNQETFCGLDLDGASSTVTFFPON 150

```
RESULT         11  
ENTRY           JN0620             #type complete  
TITLE           UDP-glucuronosyltransferase (EC 2.4.1.-) 2B-10 precursor -  
                human  
ORGANISM        #formal_name Homo sapiens #common_name man  
DATE            31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change  
                28-May-1999  
ACCESSIONS      JN0619  
REFERENCE
```

ENTRY	TITLE	ORGANISM	DATE	RESULT
12	G69998			#type complete
	thioredoxin H1 homolog ytfp -	Bacillus subtilis		
	#format_name Bacillus subtilis			
	05-Dec-1997	#sequence_revision 05-Dec-1997	#extl_change	
	11-Jun-1999			
	G69998			
	A69580			
	Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero, M.G.; Bessieres, P.; Bolotin, A.; Borchert, S.; Boriss, R.; Boursier, L.; Brans, A.; Braun, M.; Brignell, S.C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi, S.K.; Codani, J.J.; Conerton, I.F.; Cummings, N.J.; Daniel, R.A.; Denzot, F.; Devine, K.M.; Duesterhoeft, A.; Ehrlich, S.D.; Emerson, P.T.; Ehtian, K.D.; Eriington, J.; Fabret, C.; Ferrari, E.; Foulger, D.; Filtz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gallero, N.; Ghim, S.Y.; Glaser, P.; Goffeau, A.; Goldightly, E.J.; Grandi, G.; Giuseppe, G.; Guy, B.J.; Haga, K.; Halech, J.; Harwood, C.R.; Henaut, A.; Hlibert, H.; Holsapet, S.; Hosono, S.; Hulio, M.F.; Itaya, M.; Jones, L.; Joris, B.; Karamata, D.; Kasahara, Y.; Klaerr-Blanchard, M.; Klein, C.; Kobayashi, Y.; Koelster, P.; Konungstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, S.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauviel, C.; Medigue, C.; Medina, N.; Mellado, R.P.; Mizuno, M.; Moestl, D.; Nakai, S.; Noback, M.; Noone, D.; O'Reilly, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, D.; Porwollik, S.; Prescott, G.A.M.; Presecan, E.; Pulic, P.; Purnelle, B.; Rapoport, G.; Rey, M.; Reynolds, S.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, E.; Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror, S.J.; Serro, P.; Shin, B.S.; Soldo, B.; Sorokin, A.; Tacconi, E.; Takagi, T.; Takahashi, H.; Takekura, K.; Takeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terstap, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.; Jin, C.J.; Miners, J.O.; Lillywhite, K.J.; Mackenzie, P.I. Biochem. Biophys. Res. Commun. (1993) 194:496-503			
	#title	CDNA cloning and expression of two new members of the human liver UDP-glucuronosyltransferase 2B subfamily.		
	#cross-references	MUID:93326164		
	#accession	JN0620		
	##molecule_type	mRNA		
	##residues	1-528	#label	JIN
	##cross-references	GB:X63359; NID:9516149; PIDN:CAA44961.1; PID:9516150		
	##experimental_source	liver		
	##classification	#superfamily glucuronosyltransferase		
	##keywords	glycoprotein; glycosyltransferase; hexosyltransferase; transmembrane protein		
	FEATURE			
	1-23			
	24-528			
	491-508			
	66,314,481			
	SUMMARY			
	Query Match	3.6%	Score 117; DB 2; Length 528;	
	Best Local Similarity	24.7%	Pred. NO. 1.22e-02;	
	Matches	21; Conservative	24; Mismatches 36; Indels 4; Gaps 4;	
	Db	133 KRLMKLOESRFDVFDAVLPFGELLAEFNFPEV-YSH-SFSPSPSRHSGCF-IFP 189		
	Qy	266 KRFEMTLPFGQFELQIGNGVQCHQSLTEFNSTYPCQCAFNGLF-LPPIQSGPQAFS 324		
	Db	190 PSYVPVVMKSLSDQMTMERVKNNL 214		
	Qy	325 AFYFMKFLNLTSEKVSQEKVTEKM 349		

```

Vandenbol, M.; Vannier, F.; Vassarotti, A.; Viari, A.;
Wandut, R.; Wedler, E.; Weller, H.; Weitenegger, T.;
Winlers, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasunoto,
K.; Yata, K.; Yoshida, K.; Yoshikawa, H.F.; Zumbstein, E.;
Yoshikawa, H.; Danchin, A.
Nature (1997) 390:249-256
The complete genome sequence of the Gram-positive bacterium
Bacillus subtilis.

#journal
#title
The complete genome sequence of the Gram-positive bacterium
Bacillus subtilis.

#cross-references MOID:98044033
#accession
#status
G69998
    preliminary; nucleic acid sequence not shown;
    translation not shown

##molecule_type DNA
##residues GB:291119; GB:LB009126; NID:g2635411;
##cross-references GB:291119; GB:LB009126; NID:g2635411;
PIDN:CAB14962.1; PID:e1185857; PID:g2635468
#experimental_source strain 168

GENETICS
#gene yfpp
CLASSIFICATION #superfamily thioredoxin; thioredoxin homology
SUMMARY #length 107 #molecular_weight 12747 #checksum 1509

Query Match 3.5%; Score 116; DB 2; Length 107;
Best Local Similarity 34.8%; Pred. No. 1.69e-02;
Matches 16; Conservative 13; Mismatches 12; Indels 5; Gaps 4;

Db 22 MESADWCP--DCRFVEPFLPELANPEFT--YYVDKDFIDTCAG 64
      :||: || | | |||::| | |::| | |::|
Oy 295 LNTSYPCYSQCAFNGIFLPLGGDFGAFSAFYV-M-KFLNLTSE 338

RESULT 13
ENTRY T12886
TITLE conserved hypothetical protein yorA - Bacillus subtilis phage
SPAC2
ORGANISM #formal name Bacillus subtilis phage SPAC2
DATE 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change
20-Aug-1999
ACCESSIONS T12886; B69922
REFERENCE 217583
AUTHORS Lazarovic, V.; Duesterhoeft, A.; Soldo, B.; Hilbert, H.;
Maueil, C.; Karamata, D.
submitted to the EMBL Data Library, August 1997
The complete nucleotide sequence of the Bacillus subtilis
SPBac2 prophage.
#accession T12886
#status preliminary; translated from GB/EMBL/DBJ
##molecule_type DNA
##residues 1-586 ##label LAZ
#cross-references EMBL:AF020713; NID:g3025478; PID:g3025600;
PIDN:A0C13095.1

REFERENCE
#authors
Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.;
Alloni, G.; Azavedo, V.; Bettero, M.G.; Bessieres, P.;
Bolotin, A.; Borriert, S.; Borriss, R.; Boursier, L.; Brans
A.; Brun, M.; Brigelli, S.C.; Bron, S.; Brouillet, S.;
Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
Choi, S.K.; Codan, J.J.; Conneron, I.F.; Cummings, N.J.;
Daniel, R.A.; Denizot, F.; Devine, K.M.; Dueterhoeft, A.;
Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.;
Fabel, C.; Ferrari, E.; Foulger, D.; Fritz, C.; Fujita,
M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier, N.; Ghim,
S.Y.; Glaser, P.; Goffeau, A.; Golightly, E.J.; Grandi, G.;
Guiseppi, G.; Guy, B.J.; Haga, K.; Halech, J.; Harwood,
C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.;
Hullo, M.P.; Itaya, M.; Jones, L.; Joris, B.; Karamata, D.;
Kashnira, Y.; Klaerr-Blanchard, M.; Klein, C.; Kobayashi,
Y.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.;
Kurita, K.; Lapidis, A.; Iadinois, S.; Lauber, J.;
Lazarevic, V.; Lee, S.M.; Medina, N.; Mellado, R.P.; Mizuno,
M.; Moestl, D.; Nakai, S.; Noback, M.; Noone, D.; O'Reilly,
M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro,

```

[illegible]

RESULT 15
ENTRY S63501 #type fragments
TITLE apyrase (EC 3.6.1.5) - human (fragments)
ALTERNATE_NAMES ATP diphosphohydrolase
ORGANISM #formal_name Homo sapiens #common_name man
DATE 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change
13-Mar-1997
ACCESSIONS S63501
REFERENCE S63501
#authors Christoforidis, S.; Papamarcaki, T.; Galaris, D.; Kellner,
R.; Tsolas, O.
#journal Eur. J. Biochem. (1995) 234:66-74
#title Purification and properties of human placental ATP
diphosphohydrolase.
#cross-references MUID:96096723
#accession S63501
#status preliminary
#molecule_type protein
#residues 1-3;4-8;9-13;14-20;21-30;31-44 #label CHR
KEYWORDS hydrolase
SUMMARY #length 44 #checksum 6585
Query Match 3.2%; Score 104; DB 2; Length 44;
Best Local Similarity 51.3%; Pred. No. 6,75e-01;
Matches 20; Conservative 6; Mismatches 10; Indels 3; Gaps 2;
Db 9 LLGRFG-KD--YNYKKVYVSDLGALDLGASTQVTF 44
|||||:|::: : ||||| ||||| |||||
QY 148 LLGRFSQKRWFSIVPEYETNNQETFGALDLGASTQVTF 186

Search completed: Fri May 5 09:23:35 2000
Job time : 307 secs.

Dd	241	LBPCHGCGKKRVVNSDLKTKPCPKRREMTLPFOQFETIOIGNGOCHOHILEFNSTY	300
Oy	241	LRDPCHFPCKRVVNSDLYKTPTCKRRFEMTLPFOQFETIOIGNGOCHOHILEFNSTY	300
Dd	301	CPSYCAFGNGLPLPLOGDFGAFSAFYVMKPLNTSEKVSQEKYTEMMKRFCAPMWEI	360
Oy	301	CYSQCAFNGIFLPLPLOGDFGFASAFYPMKPLNTSEKVSQEKYTEMMKRFCAPMWEI	360
Dd	361	KTSYGAVKEKYLSEKCFSGTYLLSLLOGGYHTADSWEHIFIKIGSDAGWTLGWLN	420
Oy	361	KTSYGAVEKEKYLSKCFSGTYLLSLLOGGYHTADSWEHIFIKIGSDAGWTLGWMLN	420
Dd	421	LTNMIPAEOPLSPLSHT 439 	
Oy	421	LTNMIPAEOPLSPLSHT 439	
RESULT	2		
ID	PCT-US99-22955-6	STANDARD:	PRT: 454 AA.
xx	xxxxxx		
DT			
DE			
xx			
CC	Sequence 6, Application PC/TUS9922955		
CC	GENERAL INFORMATION:		
CC	APPLICANT: Maliszewski, Charles R.		
CC	APPLICANT: Gayle III, Richard B.		
CC	APPLICANT: Price, Virginia L.		
CC	APPLICANT: Gimpel, Steven D.		
CC	APPLICANT: Immunex Corporation		
CC	TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment		
CC	FILE REFERENCE: 2879-WO		
CC	CURRENT APPLICATION NUMBER: PCT/US99/22955		
CC	CURRENT FILING DATE: 1999-10-13		
CC	EARLIER APPLICATION NUMBER: US 60/104,585		
CC	EARLIER FILING DATE: 1998-10-16		
CC	EARLIER APPLICATION NUMBER: US 60/107,466		
CC	EARLIER FILING DATE: 1998-11-06		
CC	EARLIER APPLICATION NUMBER: US 60/149,010		
CC	EARLIER FILING DATE: 1999-08-13		
CC	NUMBER OF SEQ ID NOS: 31		
CC	SOFTWARE: Patentln Ver. 2.0		
CC	SEQ ID NO 6		
CC	LENGTH: 454		
CC	TYPE: PRT		
CC	ORGANISM: Artificial Sequence		
CC	SEQUENCE 454 AA; 51543 MW; 1133800 CN;		
SO			
Query Match	100.0%;	Score 3275;	DB 1; Length 454;
Best Local Similarity	100.0%;	Pred. No. 0.00e+00;	
Matches 439;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0
Dd	16	TGNKALPENVKKGIVLDGSSHTSLSYIKWPAEKENDDGVVHOVEBCEVYKPGISKPFYOK	75
Oy	1	TGNKALPENVKKGIVLDGSSHTSLSYIKWPAEKENDDGVVHOVEBCEVYKPGISKPFYOK	60
Dd	76	VNEISITLYLTDCKERAREVI PRSQHOEFYVYGATRGMRLLMESEELADRYLVYENSLS	135
Oy	61	VNEISITLYLTDCKERAREVI PRSQHOEFYVYGATRGMRLLMESEELADRYLVYENSLIS	120
Dd	136	NYPPEOGARITTGEBGAYGMITINLYLGKFSQKTRFSTVPYETNNQETFFGALDGGA	195
Oy	121	NYPPEOGARITTGEBGAYGMITINLYLGKFSQKTRNFSTVPYETNNQETFFGALDGGA	180
Dd	196	STOVYFVNQONTIESPDNALORLTGDKYNYNTHSFLLCYGKDQALMQRLADIYASNEI	255
Oy	181	STOVYFVNQONTIESPDNALORLTGKRYINYTHSFLLCYGKDQALMQRLADIYASNEI	240
Dd	256	LBDPCFHGCGKKRVVNSDLKTKPTCKRREMTLPFOQFETIOIGNGOCHOHILEFNSTY	315

Qy	241	L R D P F H G H K Y K V A N S D L Y T K P C K R R E M L I P F O Q F L O G I G N T O C H O S I L E L F N T S Y	300
Db	316	C P Y S O A E N G I F L P L O G D F G A S A F Y V M F L N I T S E K V S O E K V T E M K R F C A Q P W E I	375
Qy	301	C P Y S O C A E N G I F L P L O G D F G A S A F Y V M F L N I T S E K V S O E K V T E M K R F C A Q P W E I	360
Db	376	K T S V A G V E K Y I S E C S E G T Y I L S I L L O G Y H P T A D S W H I F I C I O S D S D G M T L G Y M L N	435
Qy	361	K T S V A G V E K Y I S E C S E G T Y I L S I L L O G Y H P T A D S W H I F I C I O S D S D G M T L G Y M L N	420
Db	436	L T N M I P A E O P L S T P L S H S T	454
Qy	421	L T N M I P A E O P L S T P L S H S T	439
RESULT	3		
ID	PCT-US99-22955-30	STANDARD;	PRT; 463 AA.
XX	AC	xxxxxxx	
XX			
DE			
XX			
CC	Sequence 30, Application PC/TUS9922955		
CC	GENERAL INFORMATION:		
CC	APPLICANT: Maliszewski, Charles R.		
CC	APPLICANT: Gayle III, Richard B.		
CC	APPLICANT: Price, Virginia L.		
CC	APPLICANT: Gimpel, Steven D.		
CC	APPLICANT: Immunex Corporation		
CC	TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment		
CC	FILE REFERENCE: 2879-WO		
CC	CURRENT APPLICATION NUMBER: PCT/US99/22955		
CC	CURRENT FILING DATE: 1999-10-13		
CC	EARLIER APPLICATION NUMBER: US 60/104,585		
CC	EARLIER FILING DATE: 1998-10-16		
CC	EARLIER APPLICATION NUMBER: US 60/107,466		
CC	EARLIER FILING DATE: 1998-11-06		
CC	EARLIER APPLICATION NUMBER: US 60/149,010		
CC	EARLIER FILING DATE: 1999-08-13		
CC	NUMBER OF SEQ ID NOS: 31		
CC	SOFTWARE: Patentln Ver. 2.0		
CC	SEQ ID NO 30		
CC	LENGTH: 463		
CC	TYPE: PRT		
CC	ORGANISM: Artificial Sequence		
CC	FEATURE:		
CC	OTHER INFORMATION: Description of Artificial Sequence: Fusion		
CC	OTHER INFORMATION: construct of human CD39		
SC	SEQUENCE 463 AA; 52621 MW; 1178782 CN;		
Query Match	100.0%;	Score 3275;	DB 1; Length 463;
Best Local Similarity	100.0%;	Pred. No. 0.00e+00;	
Matches 439;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0
Db	25	TÖNKRLPENVKYGYILDGSSHTSLXYIKKPAEKENDGVVHÖVEECRVKPGISKEVÖK	84
Qy	1	TÖNKRLPENVKYGYILDGSSHTSLXYIKKPAEKENDGVVHÖVEECRVKPGISKEVÖK	60
Db	85	VNEIGIYLTCDCMERAREVIIPRSÖHOETPVLYLGATAGMRLMESEELADRYLDVVERSLS	144
Qy	61	VNEIGIYLTCDCMERAREVIIPRSÖHOETPVLYLGATAGMRLMESEELADRYLDVVERSLS	120
Db	145	NYPDFOGARIIITGOEBEAYGMITINLYLLGFPSÖKTREFSIVPPEYNNOETFGALDLGA	204
Qy	121	NYPDFOGARIIITGOEBEAYGMITINLYLLGFPSÖKTREFSIVPPEYNNOETFGALDLGA	180
Db	205	STOVTFVPONOTIESPDNALÖFLRYTGKDYNYTTHSFLCYGKDÖALMÖKLADIÖVASNEI	264
Qy	181	STÖTFEVÖNÖTIESPNDALÖFLRYTGKDYNYTTHSFLCYGKDÖALMÖKLADIÖVASNEI	240

Db	265	LRDPFHGHGKKVNVNSLTKTPCTKREMTLPPQOEIEIGIGNYQCHSIIELFNTSY	324
Oy	241	LRDPFHGHGKKVNVNSLTKTPCTKREMTLPPQOEIEIGIGNYQCHSIIELFNTSY	300
Db	325	CPYSOCANGLFLPPLQDGFASFVFMKFLMLTSEKVSQEKVTEBMMKFCQAPBEI	384
Oy	301	CPYSOCANGLFLPPLQDGFASFVFMKFLMLTSEKVSQEKVTEBMMKFCQAPBEI	360
Db	385	KTSYAGVKEKYLSEYCEGSGVYIISLLQGYHFTADSMETHIFICKIQGSDAGWTLGYMLN	444
Oy	361	KTSYAGVKEKYLSEYCEGSGVYIISLLQGYHFTADSMETHIFICKIQGSDAGWTLGYMLN	420
Db	445	LTNMIAPQPLSTPLSHST	463
Oy	421	LTNMIAPQPLSTPLSHST	439

RESULT	4			
ID	PCT-US99-22955-27	STANDARD;	PRT:	464 AA.
XX	xxxxxx			
AC				
XX				
DT				
XX				
DE	Sequence 27, Application PC/TUS9922955			
CC	Sequence 27, Application PC/TUS9922955			
CC	GENERAL INFORMATION:			
CC	APPLICANT: Maliszewski, Charles R.			
CC	APPLICANT: Gayle III, Richard B.			
CC	APPLICANT: Price, Virginia L.			
CC	APPLICANT: Gimpel, Steven D.			
CC	APPLICANT: Immunex Corporation			
CC	TITLE OR INVENTION: Inhibitors of Platelet Activation and Recruitment			
CC	FILE REFERENCE: 2879-WO			
CC	CURRENT APPLICATION NUMBER: PCT/US99/22955			
CC	CURRENT FILING DATE: 1999-10-13			
CC	EARLIER APPLICATION NUMBER: US 60/104,585			
CC	EARLIER FILING DATE: 1998-10-16			
CC	EARLIER APPLICATION NUMBER: US 60/107,466			
CC	EARLIER FILING DATE: 1998-11-06			
CC	EARLIER APPLICATION NUMBER: US 60/149,010			
CC	EARLIER FILING DATE: 1999-08-13			
CC	NUMBER OF SEQ ID NOS: 31			
CC	SOFTWARE: PatentIn Ver. 2.0			
CC	SEQ ID NO 27			
CC	LENGTH: 464			
CC	TYPE: PRt			
CC	ORGANISM: Artificial Sequence			
CC	FEATURE:			
CC	OTHER INFORMATION: Description of Artificial Sequence: Fusion			
CC	OTHER INFORMATION: construct of human CD39			
SQ	SEQUENCE 464 AA; 52754 MW; 1182698 CN;			
Query Match 100.0%; Score 3275; DB 1; Length 464;				
Best Local Similarity 100.0%; Pred. Mismatches 0; Indels 0; Gaps 0;				
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
DB	26 TONKALPENVYGVILADGSSHTSLIYIKWPAEKENDPGVVHGYVECRVKGPGISKEPVOK 85			
OY	1 TONKALPENVYKVXGIVLDAGSSHTSLIYIKWPAEKENDPGVVHQYECEVRVKGPGISKEPVOK 60			
DB	86 VNEIGITITDCMERAREVIIPRSQHETPYVYGATAGMRLMESEBELADRYLDVVERSLS 145			
OY	61 VNEIGITITDCMERAREVIIPRSQHETPYVYGATAGMRLMESEBELADRYLDVVERSLS 120			
DB	146 NYPFPOGARIIITGEEBAYGMITNTNYLGKFSQTRMFSLVPYETUNNOETFGALDAGA 205			
OY	121 NYPFPOGARIIITGEEBAYGMITNTNYLGKFSQTRMFSLVPYETUNNOETFGALDAGA 180			
DB	206 STGVTFVPONOTIESPDNALQFRLYGKDYNYTHSFCLCYGDQALMKDLAKDIVASNEI 265			
OY	181 STGVTFVPONOTIESPDNALQFRLYGKDYNYTHSFCLCYGDQALMKDLAKDIVASNEI 240			

Db	266	LRDPCHFGYKKVNVNSDLYRTCTCKRREMTLPFQOEFIGIGNYQOCHSILELFTSY	325
Qy	241	LRDPCHFGYKKVNVNSDLYRTCTCKRREMTLPFQOEFIGIGNYQOCHSILELFTSY	300
Db	326	CPYSCAPNGFIPLPLQDPCFAFSAFYVVKF.LNLITSEKVSQEKVTEEMMKKFCAPPEEI	385
Qy	301	CPYSCAPNGFIPLPLQDPCFAFSAFYVVKF.LNLITSEKVSQEKVTEEMMKKFCAPPEEI	360
Db	386	KTSYAGVKEKYLSEYCFSGTYILSLILQGYHPTADSMEHIFIKTIGQSDAGMTLGYMLN	445
Qy	361	KTSYAGVKEKYLSEYCFSGTYILSLILQGYHPTADSMEHIFIKTIGQSDAGMTLGYMLN	420
Db	446	LTNMTAPQPLSTPLSHST	464
Qy	421	LTNMTAPQPLSTPLSHST	439

RESULT	5			
ID	PCT-US99-22955-29	STANDARD;	PRT:	473 AA.
XX AC	xxxxxx			
XX XX				
XX DT				
DE	Sequence 29, Application PC/TUS9922955			
XX				
CC	Sequence 29, Application PC/TUS9922955			
CC	GENERAL INFORMATION:			
CC	APPLICANT: Maliszewski, Charles R.			
CC	APPLICANT: Gayle III, Richard B.			
CC	APPLICANT: Price, Virginia L.			
CC	APPLICANT: Gimpel, Steven D.			
CC	APPLICANT: Immunex Corporation			
CC	TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment			
CC	FILE REFERENCE: 2879-WO			
CC	CURRENT APPLICATION NUMBER: PCT/US99/22955			
CC	CURRENT FILING DATE: 1999-10-13			
CC	EARLIER APPLICATION NUMBER: US 60/104,585			
CC	EARLIER FILING DATE: 1998-10-16			
CC	EARLIER APPLICATION NUMBER: US 60/107,466			
CC	EARLIER FILING DATE: 1998-11-06			
CC	EARLIER APPLICATION NUMBER: US 60/149,010			
CC	EARLIER FILING DATE: 1999-08-13			
CC	NUMBER OF SEQ ID NOS: 31			
CC	SOFTWARE: PatentIn Ver. 2.0			
CC	SEQ ID NO 29			
CC	LENGTH: 473			
CC	TYPE: PRT			
CC	ORGANISM: Artificial Sequence			
CC	FEATURE:			
CC	OTHER INFORMATION: Description of Artificial Sequence: Fusion			
CC	OTHER INFORMATION: construct of human CD39			
SO	SEQUENCE 473 AA; 53745 MW; 1229492 CN;			
	Query Match 100.0%; Score 3275; DB 1; Length 473;			
	Best Local Similarity 100.0%; Pred. No. 0.00e+00;			
	Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0			
Db	35 TONKALPENVKYIGVILDAAGSSHTSLIYIKWPAEKENDTGYYHGYECCRKYKPGISKRFVK 94 1 TONKALPENVKYIGVILDAAGSSHTSLIYIKWPAEKENDTGYYHGYECCRKYKPGISKRFVK 60			
Db	95 VAEIETIYLDCERARREVIIPRSQHETPVLYGATAGMRLMESEELADRYLDVVERSL 154 61 VAEIETIYLDCERARREVIIPRSQHETPVLYGATAGMRLMESEELADRYLDVVERSL 120			
Db	155 NYPFPOGARITGGEECAVGWITNTNYLLGKFSQKTRFSIVPEFTNNOETFGLADIGA 214 121 NYPFPOGARITGGEECAVGWITNTNYLLGKFSQKTRFSIVPEFTNNOETFGLADIGA 180			
Db	215 STGVTFVNQNOTIESPDNALPFLYGKDYNVYTHSFLCYGKDQALWKOKLANDIOVASNEI 274 			

QY 181 STGVTFVPONOTIESPDNALQFRLYGKDYNNYTHSFLCYGKDQALMOKLAKDIQVASNEI 240
DB 275 LRDPCHFBGKYKVVVNSDLTKPTCKRREMTLPFOQFEIQTGIGNQOCHOSILEFNTSY 334
QY 241 LRDPCHFBGKYKVVVNSDLTKPTCKRREMTLPFOQFEIQTGIGNQOCHOSILEFNTSY 300
DB 335 CPYSQCAFNGITFLPPLQDGFCAFSAFYFVMKFLNLTSEKVSQEKYTEMMKKRFGCAQPMWEI 394
QY 301 CPYSQCAFNGITFLPPLQDGFCAFSAFYFVMKFLNLTSEKVSQEKYTEMMKKRFGCAQPMWEI 360
DB 395 KTSYAGVKEKYLEXCFSGTYILSLLOGYHFTADSWEHIFIGKIQGSDAGWTLGMYLN 454
QY 361 KTSYAGVKEKYLEXCFSGTYILSLLOGYHFTADSWEHIFIGKIQGSDAGWTLGMYLN 420
DB 455 LTNMIPAEQPLSTPLSHST 473
QY 421 LTNMIPAEQPLSTPLSHST 439

RESULT 6
ID PCT-US99-22955-28 STANDARD: PRT: 474 AA.

AC xxxxxx

Sequence 28, Application PC/TUS9922955

CC GENERAL INFORMATION:
CC APPLICANT: Maliszewski, Charles R.
CC APPLICANT: Gayle III, Richard B.
CC APPLICANT: Price, Virginia L.
CC APPLICANT: Gimpe, Steven D.
CC APPLICANT: Immunex Corporation
CC TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment
CC FILE REFERENCE: 2879-WO
CC CURRENT APPLICATION NUMBER: PCT/US99/22955
CC EARLIER FILING DATE: 1999-10-13
CC EARLIER APPLICATION NUMBER: US 60/104,585
CC EARLIER FILING DATE: 1998-10-16
CC EARLIER APPLICATION NUMBER: US 60/107,466
CC EARLIER FILING DATE: 1998-11-06
CC EARLIER APPLICATION NUMBER: US 60/149,010
CC EARLIER FILING DATE: 1999-08-13
CC NUMBER OF SEQ ID NOS: 31
CC SOFTWARE: Patentln Ver. 2.0
CC SEQ ID NO 28
CC LENGTH: 474
CC TYPE: PRT
CC ORGANISM: Artificial Sequence
CC FEATURE:
CC OTHER INFORMATION: Description of Artificial Sequence: Fusion
CC OTHER INFORMATION: construct of human CD39
CC SEQUENCE 474 AA; 53816 MW; 1234400 CN;

Query Match 100.0%; Score 3275; DB 1; Length 474;

Best Local Similarity 100.0%; Pred. No. 0.00e+00; Indels 0; Gaps 0;

Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 36 TONKALPENVKYGYLDGSSHTSLYIKWPAEKENDGVVHQBECRVKPGISKFYOK 95
QY 1 TONKALPENVKYGYLDGSSHTSLYIKWPAEKENDGVVHQBECRVKPGISKFYOK 60
DB 96 VNEIGIYITLDCMERAREVIPSQHOETPVYLGATAGMRLRMESEELADRYLDVVERSL 155
QY 61 VNEIGIYITLDCMERAREVIPSQHOETPVYLGATAGMRLRMESEELADRYLDVVERSL 120
DB 156 NYPDFOGARITITGEBEAYGMITINYLKFSOKTRWFSTVPEYENNOETFGALDIGA 215
QY 121 NYPDFOGARITITGEBEAYGMITINYLKFSOKTRWFSTVPEYENNOETFGALDIGA 180
DB 216 STGVTFVPONOTIESPDNALQFRLYGKDYNNYTHSFLCYGKDQALMOKLAKDIQVASNEI 275

QY 181 STGVTFVPONOTIESPDNALQFRLYGKDYNNYTHSFLCYGKDQALMOKLAKDIQVASNEI 240
DB 276 LRDPCHFBGKYKVVVNSDLTKPTCKRREMTLPFOQFEIQTGIGNQOCHOSILEFNTSY 335
QY 241 LRDPCHFBGKYKVVVNSDLTKPTCKRREMTLPFOQFEIQTGIGNQOCHOSILEFNTSY 300
DB 336 CPYSQCAFNGITFLPPLQDGFCAFSAFYFVMKFLNLTSEKVSQEKYTEMMKKRFGCAQPMWEI 395
QY 301 CPYSQCAFNGITFLPPLQDGFCAFSAFYFVMKFLNLTSEKVSQEKYTEMMKKRFGCAQPMWEI 360
DB 396 KTSYAGVKEKYLEXCFSGTYILSLLOGYHFTADSWEHIFIGKIQGSDAGWTLGMYLN 455
QY 361 KTSYAGVKEKYLEXCFSGTYILSLLOGYHFTADSWEHIFIGKIQGSDAGWTLGMYLN 420
DB 456 LTNMIPAEQPLSTPLSHST 474
QY 421 LTNMIPAEQPLSTPLSHST 439

RESULT 7
ID PCT-US99-22955-3 STANDARD: PRT: 476 AA.

AC xxxxxx

Sequence 3, Application PC/TUS9922955

CC GENERAL INFORMATION:
CC APPLICANT: Maliszewski, Charles R.
CC APPLICANT: Gayle III, Richard B.
CC APPLICANT: Price, Virginia L.
CC APPLICANT: Gimpe, Steven D.
CC APPLICANT: Immunex Corporation
CC TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment
CC FILE REFERENCE: 2879-WO
CC CURRENT APPLICATION NUMBER: PCT/US99/22955
CC EARLIER FILING DATE: 1999-10-13
CC EARLIER APPLICATION NUMBER: US 60/104,585
CC EARLIER FILING DATE: 1998-10-16
CC EARLIER APPLICATION NUMBER: US 60/107,466
CC EARLIER FILING DATE: 1998-11-06
CC EARLIER APPLICATION NUMBER: US 60/149,010
CC EARLIER FILING DATE: 1999-08-13
CC NUMBER OF SEQ ID NOS: 31
CC SOFTWARE: Patentln Ver. 2.0
CC SEQ ID NO 3
CC LENGTH: 476
CC TYPE: PRT
CC ORGANISM: Artificial Sequence
CC FEATURE:
CC OTHER INFORMATION: Description of Artificial Sequence: Fusion
CC OTHER INFORMATION: construct of human CD39
CC SEQUENCE 476 AA; 54177 MW; 1244820 CN;

Query Match 100.0%; Score 3275; DB 1; Length 476;

Best Local Similarity 100.0%; Pred. No. 0.00e+00; Indels 0; Gaps 0;

Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 38 TONKALPENVKYGYLDGSSHTSLYIKWPAEKENDGVVHQBECRVKPGISKFYOK 97
QY 1 TONKALPENVKYGYLDGSSHTSLYIKWPAEKENDGVVHQBECRVKPGISKFYOK 60
DB 98 VNEIGIYITLDCMERAREVIPSQHOETPVYLGATAGMRLRMESEELADRYLDVVERSL 157
QY 61 VNEIGIYITLDCMERAREVIPSQHOETPVYLGATAGMRLRMESEELADRYLDVVERSL 120
DB 158 NYPDFOGARITITGEBEAYGMITINYLKFSOKTRWFSTVPEYENNOETFGALDIGA 217
QY 121 NYPDFOGARITITGEBEAYGMITINYLKFSOKTRWFSTVPEYENNOETFGALDIGA 180

Db	218	STQVFEVFNQNOTIESPDNALQRLKGRKYNVYTHSFLCYGKDQALMOKLADIOVASNEI	277
Qy	181	STQVTFVFNQNOTIESPDNALQRLKGRKYNVYTHSFLCYGKDQALMOKLADIOVASNEI	240
Db	278	LKDPCHFHGKRVKVVNDVLYKTPCTKREEMTLPFQOFETIOIGINYQOCQCHOSTLEFNYSY	337
Qy	241	LKDPCHFHGKRVKVVNDVLYKTPCTKREEMTLPFQOFETIOIGINYQOCQCHOSTLEFNYSY	300
Db	338	CPYSQCAFNGIFLPLQGDGFGAFSAFYVVMKFLNLTSEKVSQEKYTEMMKKFCAOPWEEI	397
Qy	301	CPYSQCAFNGIFLPLQGDGFGAFSAFYVVMKFLNLTSEKVSQEKYTEMMKKFCAOPWEEI	360
Db	398	KTSYVGEVEKLTSEKCFSGTYILSLLOGYHPTADMSMHIFIGTIOGSDPGMTLGYMLN	457
Qy	361	KTSYVGEVEKLTSEKCFSGTYILSLLOGYHPTADMSMHIFIGTIOGSDPGMTLGYMLN	420
Db	458	LTNMIPAEQPLSTPLSHST 476	
Qy	421	LTNMIPAEQPLSTPLSHST 439	
RESULT	8		
ID	PCT-US99-22955-8	STANDARD:	PRT: 478 AA.
XX	xxxxxx		
DT			
XX			
DE	Sequence 8, Application PC/TUS99222955		
XX			
CC	Sequence 8, Application PC/TUS99222955		
CC	GENERAL INFORMATION:		
CC	APPLICANT: Maliszewski, Charles R.		
CC	APPLICANT: Gayle III, Richard B.		
CC	APPLICANT: Price, Virginia D.		
CC	APPLICANT: Gimpel, Steven D.		
CC	APPLICANT: Immunex Corporation		
CC	TITLE OR INVENTION: Inhibitors of Platelet Activation and Recruitment		
CC	FILE REFERENCE: 2879-WO		
CC	CURRENT APPLICATION NUMBER: PCT/US99/22955		
CC	CURRENT FILING DATE: 1999-10-13		
CC	EARLIER APPLICATION NUMBER: US 60/104,585		
CC	EARLIER FILING DATE: 1998-10-16		
CC	EARLIER APPLICATION NUMBER: US 60/107,466		
CC	EARLIER FILING DATE: 1998-11-06		
CC	EARLIER APPLICATION NUMBER: US 60/149,010		
CC	EARLIER FILING DATE: 1999-08-13		
CC	NUMBER OF SEQ ID NOS: 31		
CC	SOFTWARE: Patent In Ver. 2.0		
CC	SEQ ID NO 8		
CC	LENGTH: 478		
CC	TYPE: PRT		
CC	ORGANISM: Artificial Sequence		
CC	SEQUENCE 478 AA; 5418 MM; 1255693 CN;		
Qy	Query Match	100.0%; Score 3275; DB 1; Length 478;	
Qy	Best Local Similarity 100.0%; Pred. No. 0.00e+00;		
Qy	Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0		
Db	40	TQNRKLPENVKYGIYLDGSSHTSLYIKRPAEKNPDGVVHQBVECKYKGGISKFYQK 99	
Qy	1	TQNRKLPENVKYGIYLDGSSHTSLYIKRPAEKNPDGVVHQBVECKYKGGISKFYQK 60	
Db	100	VNEIICITVITDCMERAREVYIPRSQHOETVYLGATPGMRLLNMESEFLADRVLDVVERSL 159	
Qy	61	VNEIICITVITDCMERAREVYIPRSQHOETVYLGATPGMRLLNMESEFLADRVLDVVERSL 120	
Db	160	NYPEDFOGARITITGOEBGAYGMITINYLIGKFSQKTRMFSTVYPTETNNOETFGALDLCGA 219	
Qy	121	NYPEDFOGARITITGOEBGAYGMITINYLIGKFSQKTRMFSTVYPTETNNOETFGALDLCGA 180	
Db	220	STQVTFVFNQNOTIESPDNALQRLKGRKYNVYTHSFLCYGKDQALMOKLADIOVASNEI 279	

QY	181	SVQYTFVFNQNI	ESPDNALQ	FRLLXGKQYNYNTHSFLCYGDAQMLQKLADIDVANS	240
Db	280	LADPCFHPGKKVNV	VDLYKTPCTKRP	EMTLPFQOEFIGIGNYQOCHOSILEFNTSY	339
QY	241	LADPCFHHGKKVNV	VDLYKTPCTKRP	EMTLPFQOEFIGIGNYQOCHOSILEFNTSY	300
Db	340	CPYSOCANGTFL	PLQDDFGAFSAFYVMKFLNLTSEKVSQEVTEMMKRFCAQPM	EEI	399
QY	301	CPYSOCANGTFL	PLQDDFGAFSAFYVMKFLNLTSEKVSQEVTEMMKRFCAQPM	EEI	360
Db	400	KTSYAGVEKYLSE	CEFGSGTYILSLLOGYHFTADSW	EHFFIKTIOGSDAGMTLGYMLN	459
QY	361	KTSYAGVEKYLSE	CEFGSGTYILSLLOGYHFTADSW	EHFFIKTIOGSDAGMTLGYMLN	420
Db	460	LTNMIPAEQPLST	PLSHST	478	
QY	421	LTNMIPAEQPLST	PLSHST	439	
RESULT	9		STANDARD;	PRT;	487 AA.
ID	PCNT-US99-22955-26				
XX	AC	xxxxxx			
DE	Sequence 26, Application PC/TUS9922955				
XX	Sequence 26, Application PC/TUS9922955				
CC	GENERAL INFORMATION:				
CC	APPLICANT: Maliszewski, Charles R.				
CC	APPLICANT: Gayle III, Richard B.				
CC	APPLICANT: Price, Virginia L.				
CC	APPLICANT: Gimpel, Steven D.				
CC	APPLICANT: Immunex Corporation				
CC	TITLE OR INVENTION: Inhibitors of Platelet Activation and Recruitment				
CC	FILE REFERENCE: 2879-WO				
CC	CURRENT APPLICATION NUMBER: PCT/US99/22955				
CC	CURRENT FILING DATE: 1999-10-13				
CC	EARLIER APPLICATION NUMBER: US 60/104,585				
CC	EARLIER FILING DATE: 1998-10-16				
CC	EARLIER APPLICATION NUMBER: US 60/107,466				
CC	EARLIER FILING DATE: 1998-11-06				
CC	EARLIER APPLICATION NUMBER: US 60/149,010				
CC	EARLIER FILING DATE: 1999-08-13				
CC	NUMBER OF SEQ. ID NOS: 31				
CC	SOFTWARE: Patent In Ver. 2.0				
CC	SEQ. ID NO 26				
CC	LENGTH: 487				
CC	TYPE: PRT				
CC	ORGANISM: Artificial Sequence				
CC	SEQUENCE 487 AA; 55240 MM; 1301482 CN;				
QY	Query Match	100.0%;	Score 3275;	DB 1;	Length 487;
QY	Best Local Similarity 100.0%;	Pred. No. 0.00e+00;	Mismatches 0;	Indels 0;	Gaps 0;
QY	Matches 439; Conservative				
Db	49	TQNKALPENNVAYG	YILDGSSHTSLXYIKWPAEKENDTG	VYHQBECRVKPGISKPFYQK	108
QY	1	TQNKALPENNVAYG	YILDGSSHTSLXYIKWPAEKENDTG	VYHQBECRVKPGISKPFYQK	60
Db	109	VNEIGIYVLTDCMERAREV	YIPRSQHQETFPVYLGA	TAGMRLRMESEELADRYLADVVERSL	168
QY	61	VNEIGIYVLTDCMERAREV	YIPRSQHQETFPVYLGA	TAGMRLRMESEELADRYLADVVERSL	120
Db	169	NYPDFQCARITITGOE	BAVYMTITNTYLLGKFSQKTRMF	SYVPEITNNQETFGALDLGGA	228
QY	121	NYPDFQCARITITGOE	BAVYMTITNTYLLGKFSQKTRMF	SYVPEITNNQETFGALDLGGA	180
Db	229	STQVTFVFNQNI	ESPDNALQ	FRLLXGKQYNYNTHSFLCYGDAQMLQKLADIDVANS	288
QY	181	STQVTFVFNQNI	ESPDNALQ	FRLLXGKQYNYNTHSFLCYGDAQMLQKLADIDVANS	240

Db	289	LRPCEHHGKRVVNV	SLYTKPTCKRREMTLP	PROQEEIOG	IGNYOOCHSILE	ENTSY	348
Qy	241	LRRDFEHGKRVVNV	SLYTKPTCKRREMTLP	PROQEEIOG	IGNYOOCHSILE	ENTSY	300
Db	349	CPYSQCAFGNFI	LPPLQDGFASF	AFYVVKF	FNLTJSEKYSOEKVT	EMMKPCAQWEI	408
Qy	301	CPYSQCAFGNFI	LPPLQDGFASF	AFYVVKF	FNLTJSEKYSOEKVT	EMMKPCAQWEI	360
Db	409	KTSTAQVEKYLSEY	CRSGVYILSL	LQGYHTADSN	EHIFGKIOGSD	AGMTLYGM	468
Qy	361	KTSTAQVEKYLSEY	CRSGVYILSL	LQGYHTADSN	EHIFGKIOGSD	AGMTLYGM	420
Db	469	LTNMPAEOP	LPSTPLSHT	487			
Qy	421	LTNMPAEOP	LPSTPLSHT	439			

ID	RESULT	10	STANDARD;	PRT;	510	AA.
xx	ID	US-08-930-921-1				
xx	AC	xxxxxx				
xx	DT					
xx	DE	Sequence 1, Application US/08930921B				
xx	CC	Sequence 1, Application US/08930921B				
CC	CC	GENERAL INFORMATION:				
CC	CC	APPLICANT: BEADOLIN, Adrien R.				
CC	CC	TITLE OF INVENTION: ATP-DIPHOSPHOHYDROLASES, PROCESS OF PURIFICATION				
CC	CC	TITLE OF INVENTION: THEREOF AND PROCESS OF PRODUCING THEREOF BY RECOMBINANT				
CC	CC	TITLE OF INVENTION: TECHNOLOGY				
CC	CC	FILE REFERENCE: ATP-DIPHOSPHOHYDROLASES-BEADOLIN				
CC	CC	CURRENT APPLICATION NUMBER: US/08/930,921B				
CC	CC	CURRENT FILING DATE: 1998-01-02				
CC	CC	EARLIER APPLICATION NUMBER: PCT/CA96/00223				
CC	CC	EARLIER FILING DATE: 1996-04-10				
CC	CC	NUMBER OF SEQ ID NOS: 8				
CC	CC	SOFTWARE: PatentIn Ver. 2.0				
CC	CC	SEQ ID NO 1				
CC	CC	LENGTH: 510				
CC	CC	TYPE: PR				
CC	CC	ORGANISM: unknown				
CC	CC	FEATURE:				
CC	CC	OTHER INFORMATION: Description of unknown Organism: unknown				
CC	CC	SEQUENCE 510 AA; 57964 MW; 1469467 CN;				
Query Match		100.0%; Score 3275; DB 14; Length 510;				
Best Local Similarity		100.0%; Pred. No. 0.00e+00;				
Matches 439; Conservative		0; Mismatches 0; Indels 0; Gaps 0;				
D	38	TQNKALPENVYGYIILADGSSHTSLIYIKWPAEKENDTVGVHGYEBCRVKPGISKPFYQK				97
Q	1	TQNKALPENVYGYIILADGSSHTSLIYIKWPAEKENDTVGVHGYEBCRVKPGISKPFYQK				60
D	98	VNEIGIYITLDCMEARREYIPRSQHOETPVYIGATAGMRLMSESEELADRLDVYENSLS				157
Q	61	VNEIGIYITLDCMERAREYIPRSQHOETPVYIGATAGMRLMSESEELADRLDVYENSLS				120
D	158	NYPPDFOGARITITGOEEBAYGMITINYLKGFSSQTRFSTVPEYETNNQETFFGALDLGA				217
Q	121	NYPPDFOGARITITGOEEBAYGMITINYLKGFSSQTRFSTVPEYETNNQETFFGALDLGA				180
D	218	STQVTFVPONOTIESPDNALQFRLYGKDYVNYTHSFLCYGKDAQLMOKLADIQVANSNEI				277
Q	181	STQVTFVPONOTIESPDNALQFRLYGKDYVNYTHSFLCYGKDAQLMOKLADIQVANSNEI				240
D	278	LRDQCFHGYGVKVVNVSDLYTKPCMKREMTLPPQOPEIIGISGVNOOCHOSILELFTMSY				337
Q	241	LRDQCFHGYGVKVVNVSDLYTKPCMKREMTLPPQOPEIIGISGVNOOCHOSILELFTMSY				300
D	338	CPYSQCAFNGIIFLPDQFGAFSAFYVMKFLNLISEKVSQEKVTEMMKKFCQAPWEEI				397

QY 301 CYYSCAENGITLPLPLDGFQAFSAFYVMFNLITSKYSQAEVTEMMKKFCQPPBEI 360

Db 398 KTSYAGVAREKYLSEYCFSGTYILSLLOGYHFTADSWEHIFICKIOGSDAGWTGYMLN 457

QY 361 KTSYAGVAREKYLSEYCFSGTYILSLLOGYHFTADSWEHIFICKIOGSDAGWTGYMLN 420

Db 458 LTNMTIPADOPPLSTPLSHST 476

QY 421 LTNMTIPADOPPLSTPLSHST 439

RESULT	11
ID	PCT-US99-22955-2 STANDARD; PRT; 510 AA.
AC	xxxxxx
XX	
XX	
DT	
XX	
DE	Sequence 2, Application PC/TUS9922955
CC	Sequence 2, Application PC/TUS9922955
CC	GENERAL INFORMATION:
CC	APPLICANT: Maliszewski, Charles R.
CC	APPLICANT: Gayle III, Richard B.
CC	APPLICANT: Price, Virginia L.
CC	APPLICANT: Gimpel, Steven D.
CC	APPLICANT: Immunex Corporation
CC	TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment
CC	FILE REFERENCE: 2879-WO
CC	CURRENT APPLICATION NUMBER: PCI/US99/22955
CC	CURRENT FILING DATE: 1999-10-13
CC	EARLIER APPLICATION NUMBER: US 60/104,585
CC	EARLIER FILING DATE: 1998-10-16
CC	EARLIER APPLICATION NUMBER: US 60/107,466
CC	EARLIER FILING DATE: 1998-11-06
CC	EARLIER APPLICATION NUMBER: US 60/149,010
CC	EARLIER FILING DATE: 1999-06-13
CC	NUMBER OF SEQ ID NOS: 31
CC	SOFTWARE: PatentIn Ver. 2.0
CC	SEQ ID NO 2
CC	LENGTH: 510
CC	TYPE: PRT
CC	ORGANISM: Homo sapiens
CC	SEQUENCE 510 AA; 57964 MW; 1469467 CN;
DB	Query Match 100.0%; Score 3275; DB 1; Length 510; Best Local Similarity 100.0%; Pred. No. 0.00e+00; Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
Dd	38 TONKALPENVYXGIVLDAGSSHTSLIYYKPAREKENDTGVAHQVECRVGPGISKEVQK 97 1 TQNKLAPENYVXYGIYLDAGSSHTSLIYYKPAREKENDTGVHQQVECRVGPGISKEVQK 60
Dd	98 VNEIGITVLTDCMERAREVIIPRSQHETPVYICATAGMRLLMSESELADRYLDIVERSLS 157 61 VNEIGITVLTDCMERAREVIIPRSQHETFPVYIGATAGMRLLMESEELADRVLVERSLS 120
Dd	158 NYPDFOGARIITOQBEBAYGWITINILKGFSOKTWFTSVLPETNNOTFFGLDLGGA 217 121 NYPDFOGARIIIOQBEBAYGWITINILKGFSOKTFWSTVPLETNNOEFGALLDGGGA 180
Dd	218 STQTVFVNQONTIESPDNALOFRLYGKDYNVTYSFLCYGDQAAMOKLKANDIOVASNEI 277 181 STQTVFVNQONTIESPDNALOFRLYGKDYNVTYSFLCYGRDAALMKLAKNDIOVASNEI 240
Dd	278 LRDPGFHFQYKKVVNVSDLYKTPTCKREEMTLPRQOEIEIGIGNYOOCHOSILELFNTSY 337 241 LRDPGFHFQYKKVVNVSDLYKTPTCKREEMTLPRQOEIEIGIGNYOOCHOSILELFNTSY 300
Dd	338 CPYSOCANGNIFLPLPLOGDGAFSAFTFWAKFLNLJTEKYVSQEKEYTDMARKFCQAQPWEEL 337 301 CPYSOCANGNIFLPLPLOGDGAFSAFTFWAKFLNLJTEKYVSQEKEYTDMARKFCQAQPWEEL 360

DB	398	KTSAVGEKILSECFEGYIILSLLOGYHFTADSMEHIFGKIOGSDAGWTLGYMLN	457
QY	361	KTSAVGEKILSEYCRSGIILSLLOGHFTADSMEHIFGKIOGSDAGWTLGYMLN	420
Db	458	LTMNIPAEOPILSTPLSHST	476
QY	421	LTMNIPAEOPILSTPLSHST	439
RESULT	12		
ID	US-08-701-460-2	STANDARD:	PRT; 510 AA.
XX	xxxxxx		
DE	Sequence 2, Application US/08701460		
XX			
CC	Sequence 2, Application US/08701460		
CC	GENERAL INFORMATION:		
CC	APPLICANT: MARCUS, AARON J.		
CC	APPLICANT: MALISZEWSKI, CHARLES R.		
CC	APPLICANT: GAYLE, RICHARD B.		
CC	TITLE OF INVENTION: METHODS OF REGULATING HEMOSTASIS AND		
CC	TITLE OF INVENTION: THROMBOSIS, AND COMPOUNDS USEFUL THEREFOR		
CC	NUMBER OF SEQUENCES: 16		
CC	CORRESPONDENCE ADDRESSES:		
CC	ADDRESSEE: Immunex Corporation		
CC	STREET: 51 University Street		
CC	City: Seattle		
CC	STATE: WA		
CC	COUNTRY: USA		
CC	ZIP: 98101		
CC	COMPUTER READABLE FORM:		
CC	MEDIUM TYPE: Floppy disk		
CC	COMPUTER: Apple Power Macintosh 7200/90		
CC	OPERATING SYSTEM: Apple Operating System 7.5.3		
CC	SOFTWARE: Microsoft Word 6.0.1 for Power Macintosh		
CC	CURRENT APPLICATION DATA:		
CC	APPLICATION NUMBER: US/08/701,460		
CC	FILING DATE: 22 AUGUST 1996		
CC	CLASSIFICATION: 514		
CC	ATTORNEY/AGENT INFORMATION:		
CC	NAME: Perkins, Patricia Anne		
CC	REGISTRATION NUMBER: 34,693		
CC	REFERENCE/DOCKET NUMBER: 2808		
CC	TELECOMMUNICATION INFORMATION:		
CC	TELEPHONE: 206-587-0430		
CC	INFORMATION FOR SEQ ID NO: 2:		
CC	SEQUENCE CHARACTERISTICS:		
CC	LENGTH: 510 amino acids		
CC	TYPE: amino acid		
CC	TOPOLOGY: linear		
CC	MOLECULE TYPE: protein		
CC	SEQUENCE 510 AA; 57964 MW; 1469467 CN;		
QY	Query Match	100.0%; Score 3275; DB 12; Length 510;	
Db	Best Local Similarity 100.0%; Pred. No. 0.00e+00;		
QY	Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Db	38 TQNALPENNVKYGIVLDAGSSHTSLYIKWPAEKENDTGVVHGYECCRYVGPISKEFVK	97	
QY	1 TQNALPENNVKYGIVLDAGSSHTSLYIKWPAEKENDTGVVHGYECCRYVGPISKEFVK	60	
Db	98 VNEIGIYLTDCMEAREVYIPRSQHQETPVYLGATAGRLLRMSESEELADRYLDVERSLS	157	
QY	61 VNEIGIYLTDCMEAREVYIPRSQHQETPVYLGATAGRLLRMSESEELADRYLDVERSLS	120	
Db	158 NYPDPGQARITLQOEAGAGWITINTLLKFSQKTRWESIYPTNNQTFGALDYGGA	217	
QY	121 NYPDPGQARITLQOEAGAGWITINTLLKFSQKTRWESIYPTNNQTFGALDYGGA	180	

D	b	218	STOYTFVPOQNTTISPDNALQFRLYGKDYNYTHSFLCYGDKDALMOKLAKDIQVASNEI	277
O	y	181	STOYTFVPOQNTTISPDNALQFRLYGKDYNYTHSFLCYGDKDALMOKLAKDIQVASNEI	240
D	b	278	LRDPCFHFGYKKVVNVSDLYKTPTCKRFEMTMLPFOQFEIOIGIGNYOQCHOSILELFNTSY	337
O	y	241	LRDPCFHFGYKKVVNVSDLYKTPTCKRFEMTMLPFOQFEIOIGIGNYOQCHOSILELFNTSY	300
D	b	338	CPSYCACAFNGIFLPPLQDGFAGFSAPFYFMVKFMLNLTSEKVSQEKVTMMKKPCAQPWEET	397
O	y	301	CPSYCACAFNGIFLPPLQDGFAGFSAPFYFMVKFMLNLTSEKVSQEKVTMMKKPCAQPWEET	360
D	b	398	KTSYAGVEKEKLSTCYSCSGTYILSLLLQGYNFTADSNHEHIFGIKGLOGSAGMTLGMYLN	457
O	y	361	KTSYAGVEKEKLSTCYSCSGTYILSLLLQGYNFTADSNHEHIFGIKGLOGSAGMTLGMYLN	420
D	b	458	LTNMIPAEOPLSTPLSHST	476
O	y	421	LTNMIPAEOPLSTPLSHST	439
R	E	S	RESULT	13
I	D	ID	US-09-374-S86-1	STANDARD;
X	X	XX		PRT; 510 AA.
X	X	XXXXX		
D	T	DT		
X	X	Sequence 1, Application US/09374586		
CC	CC	Sequence 1, Application US/09374586		
CC	CC	GENERAL INFORMATION:		
CC	CC	APPLICANT: Plinsky, David J.		
CC	CC	TITLE OF INVENTION: CD39/ECTO-ADPASE AS A TREATMENT FOR THROMBOTIC AND		
CC	CC	TITLE OF INVENTION: ISCHEMIC DISORDERS		
CC	CC	FILE REFERENCE: 0575/59167		
CC	CC	CURRENT APPLICATION NUMBER: US/09/374, 586		
CC	CC	CURRENT FILING DATE: 1999-08-13		
CC	CC	NUMBER OF SEQ ID NOS: 2		
CC	CC	SOFTWARE: PatentIn Ver. 2.1		
CC	CC	SEQ ID NO 1		
CC	CC	LENGTH: 510		
CC	CC	TYPE: PRT		
CC	CC	ORGANISM: HOMO-SAPIEN		
S	Q	SEQUENCE	510 AA; 57964 MW; 1469467 CN;	
		Query Match	100.0%; Score 3275; DB 25; Length 510;	
		Best Local Similarity	100.0%; Pred. No. 0.00e+00;	
		Matches	439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
D	b	38	TQNALPENNVKYGIVLDAGSSHTSLYTKPKPAEKENDTGVHVQVEBCRVKGPISKRFVOK	97
O	y	1	TQNALPENNVKYGIVLDAGSSHTSLYTKPKPAEKENDTGVHVQVEBCRVKGPISKRFVOK	60
D	b	98	VNEIGIYITDCMEAREVIIPRSQHQTPEVYLGAAGRLLRMESEELADRYLDIVERSLS	157
O	y	61	VNEIGIYITDCMEAREVIIPRSQHQTPEVYLGAAGRLLRMESEELADRYLDIVERSLS	120
D	b	158	NYPDPFGARIIITOEGSAVGMITINLLKFSQKTWESIVELYENNNOETFGALDLGA	217
O	y	121	NYPDPFGARIIITOEGSAVGMITINLLKFSQKTWESIVELYETTNETFGALDLGA	180
D	b	218	STOYTFVPONOTISSPNALQFRLYGKDYNYTHSFLCYGDKDALMOKLAKDIQVASNEI	277
O	y	181	STOYTFVPONOTISSPNALQFRLYGKDYNYTHSFLCYGDKDALMOKLAKDIQVASNEI	240
D	b	278	LRDPCFHFGYKKVVNVSDLYKTPTCKRFEMTMLPFOQFEIOIGIGNYOQCHOSILELFNTSY	337
O	y	241	LRDPCFHFGYKKVVNVSDLYKTPTCKRFEMTMLPFOQFEIOIGIGNYOQCHOSILELFNTSY	300
D	b	338	CPSYCACAFNGIFLPPLQDGFAGFSAPFYFMVKFMLNLTSEKVSQEKVTMMKKPCAQPWEET	397
O	y	301	CPSYCACAFNGIFLPPLQDGFAGFSAPFYFMVKFMLNLTSEKVSQEKVTMMKKPCAQPWEET	360

D	b		398	KTSVAGVEKYLSECEGTYILSLLOGCYHPTDMSWHIFPKIGQSDSGMPLGLWLN	457
O	y		361	KTSVAGVEKYLSECEGTYILSLLOGCYHPTADSWEHIFPKIGQSDSGMPLGLWLN	420
D	b		458	LTMNIPAEOPLPSTPLSHST	476
O	y		421	LTMNIPAEOPLPSTPLSHST	439
		RESULT	14		
ID		PCT-US99-22955-4	STANDARD:	PRT;	476 AA.
AC		xxxxxx			
XX					
XX					
De					
Sequence	4,	Application PC/TUS9922955			
CC					
CC		Sequence 4, Application PC/TUS9922955			
CC		GENERAL INFORMATION:			
CC		APPLICANT: Maliszewski, Charles R.			
CC		APPLICANT: Gayle III, Richard B.			
CC		APPLICANT: Price, Virginia L.			
CC		APPLICANT: Gimpel, Steven D.			
CC		APPLICANT: Immunex Corporation			
CC		TITLE OF INVENTION: Inhibitors of Platelet Activation and Recruitment			
CC		FILE REFERENCE: 2879-WO			
CC		CURRENT APPLICATION NUMBER: PCT/US99/22955			
CC		CURRENT FILING DATE: 1999-10-13			
CC		EARLIER APPLICATION NUMBER: US 60/104,585			
CC		EARLIER FILING DATE: 1998-10-16			
CC		EARLIER APPLICATION NUMBER: US 60/107,466			
CC		EARLIER FILING DATE: 1998-11-06			
CC		EARLIER APPLICATION NUMBER: US 60/149,010			
CC		EARLIER FILING DATE: 1999-08-13			
CC		NUMBER OF SEQ ID NOS: 31			
CC		SOFTWARE: PatentIn Ver. 2.0			
CC		SEQ ID NO 4			
CC		LENGTH: 476			
CC		TYPE: PRT			
CC		ORGANISM: Artificial Sequence			
CC		FEATURE:			
CC		OTHER INFORMATION: Description of Artificial Sequence: Fusion			
CC		OTHER INFORMATION: construct of human CD39			
CC		FEATURE:			
CC		NAME/KEY: VARIANT			
CC		LOCATION: (39)			
CC		OTHER INFORMATION: Any amino acid, preferably Cys or Ser			
SQ		SEQUENCE	476 AA; 54078 MW; 1246383 CN;		
		Query Match	97.9%; Score 3206; DB 1; Length 476;		
		Best Local Similarity 100.0%; Pred. No. 0.00e+00;			
		Matches 428; Conservativity 0; Mismatches 0; Indels 0; Gaps 0			
D	b		49	YGIIVDAGSSHTSLIYKMPAEEKNDTGVAQVEECIRKFGISKFYQKVMEIGIYITDC	108
O	y		12	YGIIVDASSHSLSLIYKMPAEEKNDTGVAQVEECIRKFGISKFYQKVMEIGIYITDC	71
D	b		109	MERAREVIPRSOHETPVYLGATAGMARLLRMESEELADRVLDDVERSLSNYPEDFQGARI	168
O	y		72	MERAREVIPRSOHETPVYLGATAGMARLLRMESEELADRVLDDVERSLSNYPEDFQGARI	131
D	b		169	ITGOEGEAVGYMTTINYYLKGFSQKTRMFSIYPIETNNDETGGALDLDGASQVYFPQNO	228
O	y		132	ITGOEGEAVGYMTTINYYLKGFSQKTRMFSIYPIETNNDETGGALDLDGASQVYFPQNO	191
D	b		229	TIESPDNALQFRILYKCDYNVTYTHSELFCYGKDQALMQKLAKDIQVASNEILLDPCHFPGYK	288
O	y		192	TIESPDNALQFRILYKCDYNVTYTHSELFCYGKDQALMQKLAKDIQVASNEILLDPCHFPGYK	251
D	b		289	KVVNVSDLYKTPCKRFEMTLRPQQFEIQTGIGNYQQCHQSILELDFNTSYCPYSOCAPENG	348

```

0Y 252 KVVNWSDLKTPCTRFEKMTLPQOFEIQLGNGVQOCHQSHLLEFNNTSYCPYSQAFNGI 311
Db 349 FLPLPQGGFGAFSAFYFWKFLNLTSEKVSQEKYTEMMKKFKCAOPWMEIKTSYAGVREKY 408
0Y 312 FLPLPQGGFGAFSAFYFWKFLNLTSEKVSQEKYTEMMKKFKCAOPWMEIKTSYAGVREKY 371
Db 409 LSEYFSGTYIISLLQGYHFTADSWEHHTIGKIQSDAGWTGYNLNTNMPAOPPL 468
0Y 372 LSEYFSGTYIISLLQGYHFTADSWEHHTIGKIQSDAGWTGYNLNTNMPAOPPL 431
Db 469 STPLSHST 476
0Y 432 STPLSHST 439

RESULT 15
ID US-08-701-460-4 STANDARD; PRT; 510 AA.
XX
XX
XX
XX
XX
DE Sequence 4, Application US/08701460
XX
CC Sequence 4, Application US/08701460
CC GENERAL INFORMATION:
CC APPLICANT: MARCUS, AARON J.
CC APPLICANT: MALISZEMSKI, CHARLES R.
CC APPLICANT: GAYLE, RICHARD B.
CC TITLE OF INVENTION: METHODS OF REGULATING HEMOSTASIS AND
CC TITLE OF INVENTION: THROMBOSIS, AND COMPOUNDS USEFUL THEREFOR
CC NUMBER OF SEQUENCES: 16
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: Immunex Corporation
CC STREET: 51 University Street
CC CITY: Seattle
CC STATE: WA
CC COUNTRY: USA
CC ZIP: 98101
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: Apple Power Macintosh 7200/90
CC OPERATING SYSTEM: Apple Operating System 7.5.3
CC SOFTWARE: Microsoft Word 6.0.1 for Power Macintosh
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/701,460
CC FILING DATE: 22 AUGUST 1996
CC CLASSIFICATION: 514
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Perkins, Patricia Anne
CC REGISTRATION NUMBER: 34,693
CC REFERENCE/DOCKET NUMBER: 2808
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 206-587-0430
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 510 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 510 AA; 57205 MW; 1483126 CN;

Query Match 80.8%; Score 2645; DB 12; Length 510;
Best Local Similarity 76.6%; Pred. No. 4,51e-280;
Matches 339; Conservative 57; Mismatches 42; Indels 4; Gaps 4;

Db 38 TQNKRLPENVXKGIYLDAGSSHTNLTIYKPADEKENDTGVVQOLECQVKGPGISKYAK 97
0Y 1 TQNKRLPENVXKGIYLDAGSSHTSLYIKPADEKENDTGVVHVEECRVKVGPGISKFYAK 60

Db 98 TDEIATVLAECMELSTELPTSKHHQTPWYIGATAGNRILRMESEQADRYLAAVTSLSK 157
0Y 61 VEIGIYLTDMERARVEIPRSQDEPTVYIGATAGNRILRMESEELADRYLDVVERSLS 120






```

```

Db 158 SYRPFQAKITIGQERGAIVGWTITINVLGRFTQEOQSWLSLIS-DSOKOETFGALDYGGA 216
QY 121 NYPDFQAKITIGQERGAIVGWTITINVLGRFSOKTRWFSIVPEITNNQETFGALDYGGA 180
Db 217 STQITFVPQNSTIESPENSIOFRLYGEDYVYTHSFICYGKDQALMOKLAKDIQVSSGGV 276
QY 181 STQITFVPQNOTIESPDNALQFRLYGKDYNVYTHSFICYGKDQALMOKLAKDIQVASSNEI 240
Db 277 LKDPCEMPGEKVVVNSSELYGTCTKREKKLPDQFRIOGTGDYEOCHOSILEFNNSH 336
QY 241 LKDPCEMPGEKVVVNSSELYGTCTKREKKLPDQFRIOGTGDYEOCHOSILEFNNSH 300
Db 337 CPYSQCAFNGVFLPPLHSGAFSAFVWDFEKVAKNSVISOEKMTEITKNPCKSWE 396
QY 301 CPYSQCAFNGVFLPPLHSGAFSAFVWDFEKVAKNSVISOEKMTEITKNPCKSWE 358
Db 397 ERTSYPSVKEKYLEXCFSGAYILSL-OGYNFTDSSWEQIHFMGRIKDSNAGWTLGYM 455
QY 359 ERTSYPSVKEKYLEXCFSGAYILSL-OGYNFTDSSWEQIHFMGRIKDSNAGWTLGYM 418
Db 456 LNTNMIPAEQPLSPPLPHST 476
QY 419 LNTNMIPAEQPLSPPLPHST 439

```

Search completed: Fri May 5 09:50:43 2000
 Job time : 835 secs.

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; S73813; AAB32152.1; -
 DR EMBL; U87967; AAB47572.1; -
 DR MIM; 601752; -
 DR PROSITE; PS01238; GDAL_CD39_NTPASE; 1.

DR PFAM; PF01150; GDAL_CD39; 1.
 KW Hydrolyase; Transmembrane; Antigen; Glycoprotein.
 FT DOMAIN 1 16
 FT TRANSMEM 17 38
 FT DOMAIN 39 478
 FT TRANSMEM 479 499
 FT DOMAIN 500 510
 FT CARBOHYD 73 73
 FT CARBOHYD 227 227
 FT CARBOHYD 292 292
 FT CARBOHYD 334 334
 FT CARBOHYD 371 371
 FT CARBOHYD 457 457
 SQ SEQUENCE 510 AA; 57964 MW; E403B5C9 CRC32;

Query Match 100.0%; Score 3275; DB 1; Length 510;
 Best Local Similarity 100.0%; Pred. No. 0.00e+00;
 Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 38 TONKALPENVKYGIYLDAGSSHTSLYIKWPAEKENDGVVHVEECRVKGPISKFVOK 97
 QY 1 TONKALPENVKYGIYLDAGSSHTSLYIKWPAEKENDGVVHVEECRVKGPISKFVOK 60
 Db 98 VNEIGIYITDCMERAREVYPRSOHOETPYLGATAGMRLRMESELDRLVLYVERSLIS 157
 QY 61 VNEIGIYITDCMERAREVYPRSOHOETPYLGATAGMRLRMESELDRLVLYVERSLIS 120
 Db 158 NYPFDFOGARITGOEGAYGWTITNYLLGKFSOKTRMFESIYPTENNOETFGALDLGGA 217
 QY 121 NYPFDFOGARITGOEGAYGWTITNYLLGKFSOKTRMFESIYPTENNOETFGALDLGGA 180
 Db 218 STQVTFVPONOTIESPDNALQFRLYGKDVNYTHSFLCYGKDQALMOKLADIOVASNEI 277
 QY 181 STQVTFVPONOTIESPDNALQFRLYGKDVNYTHSFLCYGKDQALMOKLADIOVASNEI 240
 Db 278 LRDPCHFGYKRVVNVSDLYKTPCTKRREPMTLRPOQFELIOGNGNQCHOSILELFTNSY 337
 QY 241 LRDPCHFGYKRVVNVSDLYKTPCTKRREPMTLRPOQFELIOGNGNQCHOSILELFTNSY 300
 Db 338 CPYSQCAENGFLPPLQDGFAGFAFYFMKFLNLTSEKVSQEKYTEMMKKFCAQPMEEI 397
 QY 301 CPYSQCAENGFLPPLQDGFAGFAFYFMKFLNLTSEKVSQEKYTEMMKKFCAQPMEEI 360
 Db 398 KTSYAGVREKYLSEYCFSGTYILSLLOGYHFTADSWEHIFHFGIKIGSDAGWTILGYMLN 457
 QY 361 KTSYAGVREKYLSEYCFSGTYILSLLOGYHFTADSWEHIFHFGIKIGSDAGWTILGYMLN 420
 Db 458 LTNMIPAEQPLSTPLSHST 476
 QY 421 LTNMIPAEQPLSTPLSHST 439

RESULT 2 STANDARD: PRT: 510 AA.
 AC P55772;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE CYTOPLASMIC (POTENTIAL).
 DE ACTIVATION ANTIGEN (CD39 ANTIGEN).
 GN CD39.
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 95015846.
 RA MALISZEWSKI C.R., DELESPESSÉ G.J.T., SCHOENBORN M.A., ARMITAGE R.J.,
 RA FANSLAW W.C., NAKAJIMA T., BAKER E., SUTHERLAND G.R., POINDexter R.,
 RA BIRKS C., ALPERT A., FRIEND D., GIMPEL S.D., GAYLE R.B. III;
 RT "The CD39 lymphoid cell activation antigen. Molecular cloning and
 RT structural characterization."
 RL J. Immunol. 153:3574-3583(1994).

CC -1- CATALYTIC ACTIVITY: ATP + 2 H(2)O -> AMP + 2 PHOSPHATE.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AF037366; AAB92259.1; -
 DR MGD; MGI:102805; CD39.
 DR PROSITE; PS01238; GDAL_CD39_NTPASE; 1.
 DR PFAM; PF01150; GDAL_CD39; 1.
 KW Hydrolyase; Transmembrane; Antigen; Glycoprotein.
 FT DOMAIN 1 16
 FT TRANSMEM 17 38
 FT DOMAIN 39 478
 FT TRANSMEM 479 499
 FT DOMAIN 500 510
 FT CARBOHYD 73 73
 FT CARBOHYD 226 226
 FT CARBOHYD 291 291
 FT CARBOHYD 333 333
 FT CARBOHYD 428 428
 FT CARBOHYD 457 457
 SQ SEQUENCE 510 AA; 57205 MW; 0570BBFE CRC32;

Query Match 80.8%; Score 2645; DB 1; Length 510;
 Best Local Similarity 76.6%; Pred. No. 0.00e+00;
 Matches 338; Conservative 57; Mismatches 42; Indels 4; Gaps 4;

Db 38 TONKALPENVKYGIYLDAGSSHTSLYIKWPAEKENDGVVHVEECRVKGPISKFVOK 97
 QY 1 TONKALPENVKYGIYLDAGSSHTSLYIKWPAEKENDGVVHVEECRVKGPISKFVOK 60
 Db 98 TDEIGAYLAECMELSTELIPTSKHQTPLYLGATAGMRLRMESELDRLVLYVERSLIS 157
 QY 61 VNEIGIYITDCMERAREVYPRSOHOETPYLGATAGMRLRMESELDRLVLYVERSLIS 120
 Db 158 NYPFDFOGARITGOEGAYGWTITNYLLGKFSOKTRMFESIYPTENNOETFGALDLGGA 216
 QY 121 NYPFDFOGARITGOEGAYGWTITNYLLGKFSOKTRMFESIYPTENNOETFGALDLGGA 180
 Db 217 STQVTFVPONOTIESPDNALQFRLYGKDVNYTHSFLCYGKDQALMOKLADIOVASNEI 276
 QY 181 STQVTFVPONOTIESPDNALQFRLYGKDVNYTHSFLCYGKDQALMOKLADIOVASNEI 240
 Db 277 LRDPCHFGYKRVVNVSDLYKTPCTKRREPMTLRPOQFELIOGNGNQCHOSILELFTNSH 336
 QY 241 LRDPCHFGYKRVVNVSDLYKTPCTKRREPMTLRPOQFELIOGNGNQCHOSILELFTNSH 300
 Db 337 CPYSQCAENGFLPPLQDGFAGFAFYFMKFLNLTSEKVSQEKYTEMMKKFCAQPMEEI 396
 QY 301 CPYSQCAENGFLPPLQDGFAGFAFYFMKFLNLTSEKVSQEKYTEMMKKFCAQPMEEI 358
 Db 397 ERTSTYAGVREKYLSEYCFSGTYILSLLOGYHFTADSWEHIFHFGIKIGSDAGWTILGYM 455
 QY 359 ERTSTYAGVREKYLSEYCFSGTYILSLLOGYHFTADSWEHIFHFGIKIGSDAGWTILGYM 418

RESULT	5	STANDARD:	PRT:	485 AA.
ID	YV4E_CAEEL			
AC	Q18411;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	HYPOTHETICAL 54.3 KD PROTEIN C33H5.14 IN CHROMOSOME IV.			
GN	C33H5.14.			
OS	Caenorhabditis elegans.			
OC	Eukaryota; Metazoa; Nematoda; Secernentea; Rhabdita; Rhabditidae;			
OC	Rhabditina; Rhabditoidae; Rhabditidae; Peloderinae; Caenorhabditis.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-BRISTOL N2.			
RA	BRADSHAW H., STELLYES L.;			
RL	Submitted (Dec-1995) to the EMBL/Genbank/DBJ databases.			
CC	-1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-slb.ch/announce/			
CC	or send an email to license@isb-slb.ch).			
CC	-----			
DR	EMBL; U41007; AAA82272.1; -.			
DR	WORMPEP; C33H5.14; CE04157.			
DR	PROSITE; PS01238; GDAL CD39_NTPASE; 1.			
DR	PFAM; PF01150; GDAL CD39; 1.			
KM	Hypothetical protein; Transmembrane; Hydrolase.			
FT	TRANSMEM 439 459 POTENTIAL.			
SQ	SEQUENCE 485 AA; 54309 MW; 00659F2B CRC32;			
Query Match	9.9%; Score 323; DB 1; Length 485;			
Best Local Similarity	25.8%; Pred. No. 1,21e-44;			
Matches	109; Conservaive 99; Mismatches 179; Indels 35; Gaps 29;			
Db	21 NNKYGVIDCAGSSGTRLFVYTLKPLSGGLTMDTLHSESPYVKVPGLSFGDKPRQ 80			
Qy	8 ENMKGYIVADASSHSHTLYIKW-P-AEKENDTG-VVHQVEECRVK-GPGISKFQKVN 63			
Db	81 VVEVYTLPLRFADENHPYQGLTETDILLFATAGMRLLPFAQDAITKLNQNGKSTLAR 140			
Qy	64 IITVLDCKERRAREVYPRSOHOETPYLPATAGMRLLRSESELADRYLDVVERSTSNP 123			
Db	141 VSDSNRIITDAGMEGYSIAVNYILGRPKENDS-KVGMIMDGASVOIAEIANE-KE 198			
Qy	124 PFGQGARITIGDEGAYGMITITINILGKRSQATRWMSIYPYETNNQETFGALDGGASQ 183			
Db	199 -SYNGN-VYEINLGSIETNEDY-K-YKISTYTFIAGYGANEGI-KYENS-L-VKSGNS-N 251			
Qy	184 VTFEVPQNTIESPDNALQFRL-YGKQYNYVTHSFLLCYGKDAQLMOKLADIIVASNEIIR 242			
Db	222 DSCSRGRLURLL-GE-FIVNGTGENDVOLA-OVSSL--IGD-KA-QPSC-P-NPT-CF 299			
Qy	243 DCEPHRGYKVVNVSQLYTPTCKRPEMLPPOQFEIIGTIGQOCHOSILELFTNSTYCP 302			
Db	300 LKNVIAVSNLSTVOL-YG-FSEYWTTSNFSGGEYHRO-KFTDEVRARYCQKDWNDIOD 356			
Qy	303 YSQCAPNGFIFLPLOGDPFASAFYVYKMFNLITSKYSQEKYTEMMKFFCAQPMEEI-- 360			
Db	357 GFKRNEFPNADIERLTGNCFKAAWTSVLADGFN-VDKTKHLFQSVLTAGEEMQALCA 415			
Qy	361 --KTS-YAGVKEKYLYSEYFFSGTYLITLLQGHYFADSEWHIHFYTKIGSDAGWTLCY 417			
Db	416 ML 417			
Qy	418 ML 419			

[illegible]

[illegible]

```

DN 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE GDANOSINE-DIPHOSPHATASE (EC 3.6.1.42) (GDPase).
GN GDA1 OR YEL042W OR SYGP-ORP16.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
   Saccharomycetaceae; Saccharomyces.
NN [1]
NN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP STRAIN-G2-9;
RC MEDLINE: 93308137.
RA ABELSON C., YAMAGISAWA K., MADON E.C., HAESTLER A., MOREMEN K.,
RA HITSCHBERG C.B., ROBBINS P.W.;
RT "Guanosine diphosphatase is required for protein and sphingolipid
RT glycosylation in the Golgi lumen of Saccharomyces cerevisiae.";
RL J. Cell Biol. 122:307-323(1993).
RN [2]
RN SEQUENCE FROM N.A.
RP MULLIGAN J.T., DIETRICH F.S., HENNESSEY K.M., SEHL P., KOMP C.,
RA WEI Y., TAYLOR P., NAKAHARA K., ROBERTS D., DAVIS R.W.;
RL Submitted (FEB-1993) to the EMBL/Genbank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RP STRAIN-S288C / AB972;
RC DIETRICH F.S., MULLIGAN J.T., HENNESSEY K.M., ALLEN E., ARAUJO R.,
RA AVILES E., BENO A., BRENNAN T., CARPENTER J., CHEN E., CHERRY J.M.,
RA CHONG E., DUNCAN M., GUZMAN E., HARTZEL G., HUNICKE-SMITH S.,
RA HWANG R., KAYSER A., KOMP C., LASKHARI D., LEW H., LIN D.,
RA MOSEDALE D., NAKAHARA K., NAMATH A., NONGREN R., OERNER P., OH C.,
RA PETER F.X., ROBERTS D., SEHL P., SCHRAM S., SHOKREN T., SMITH V.,
RA TAYLOR P., WEI Y., YELTON M., BOSTEIN D., DAVIS R.W.;
RL Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: AFTER TRANSFER OF SUGARS TO ENDOGENOUS MACROMOLECULAR
CC ACCEPTORS, THE ENZYME CONVERTS NUCLEOSIDE DIPHOSPHATES TO
CC NUCLEOSIDE MONOPHOSPHATES WHICH IN TURN EXIT THE GOLGI LUMEN IN
CC A COUPLED ANTIPORTER REACTION, ALLOWING ENTRY OF ADDITIONAL
CC NUCLEOTIDE SUGAR FROM THE CYTOSOL.
CC -1- CATALYTIC ACTIVITY: GDP + H(2)O = GMP + ORTHOPHOSPHATE.
CC -1- PATHWAY: GLYCOSYLATION.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN, GOLGI.
CC -1- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, L19560; AAA34656.1; -
DR EMBL, U18779; AAB65000.1; -
DR PIR, S30837; S30837.
DR PIR, A40732; A40732.
DR SGD, L0000695; GDA1.
DR PROSITE: PS01238; GDA1_CD39_NTPASE; 1.
DR PFAM: PF01150; GDA1_CD39; 1.
KW Hydrolase; Golgi stack; Glycoprotein; Transmembrane; Signal-anchor.
FT DOMAIN 1 9 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 10 24 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT DOMAIN 25 518 LUMENAL (POTENTIAL).
FT CARBOHYD 41 41 POTENTIAL.
FT CARBOHYD 280 280 POTENTIAL.
FT CARBOHYD 335 335 POTENTIAL.
SO SEQUENCE 518 AA; 56821 MW; D139338C CRC32;
Query Match 8.1%; Score 265; DB 1; Length 518;
Best Local Similarity 27.8%; Pred. No. 5,76e-32;
Matches 122; Conservative 92; Mismatches 180; Indels 45; Gaps 37;

```

```

QY 7 PENNYGIIVLDAGSSHTSLIYKMPAEKENDTGVVHVECEVRKGPISKF-VQKNEIG 65
DB 145 S-LDPLLVAMNYVPKARSCPTPAVKATAGLRLLGDASKSKILSAVRDLEKDP-FPV 202
QY 66 IYLDGCMARAREVIRSRQHEPVLGATAGMRLL-RMSEELADRLDYVERSLNTPF 124
DB 203 VEGDGVSIMGDDEGEVFAWITTYLLGNIGANGP--KL-P--TAA--VF--DLGGSGSTQ 252
QY 125 -DFGAGRIITGEEGAYGMITTYLLGKFSQKTRMFSIYEPYETNNQETGALDGGASTQ 183
DB 253 IVFETFPINEMKVDGEKFDLKGEDENTLYQFSLHGLGKEG-RNKV-NSVLVE-NA- 308
QY 184 VTFPQONOTIS-PNALQFRL-YKGD-YNVYTHSFLCGKQOALMOKLAKOIQVANSNI 240
DB 309 LKDGKILKDMTKTHTQLSSCPDPRKYNATNEKVTLESKETYITIDFGPEPSGACQCFIT 368
QY 241 LRDP-CFHPGYKRYVNVVDLYTP-CTKRF-MTLFPQO-FEIOGIGNQOC-HQS-IL- 293
DB 369 DELINKDAOCSPSCFENGWHQPSLVTRTFKESNDIYIFSIFYDRTRPLGMLPSTLNEIN 428
QY 294 -ELEFMT-SYCPYSQCAFNGIIFLPLOGDF--GA-FSAP-YFVWKFLNL-LSEKVSQEKVT 346
DB 429 DLARVCGEETWNSVFSGIAGSLDELSDFCLDSFQVSLHTGYIPIQ-RE-LKT 486
QY 347 EMMKRF-C-AQ-PWEIKIKSYAGVKEKRYLSE-YCFSGTYILSLLOGYHFTADSWEHIF 402
DB 487 GKTIANKKEIGWCLGASLP 505
QY 403 IGTGSGDAGWTLGTMNL 421

RESULT 9
ID NTP2_TOXGO STANDARD; PRT; 628 AA.
AC 027895: 027798: 027801:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE NUCLEOSIDE-TRIPHOSPHATASE II PRECURSOR (EC 3.6.1.15) (NTPASE-II)
GN NTP1.
OS Toxoplasma gondii.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Sarcocystidae;
OC Toxoplasma.
RN [1]
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.
RC STRAIN-RH, AND BEVERLEY;
RX MEDLINE; 95263459.
RA ASAI T., MIDRA S., SIBLEY L.D., OKABAYASHI H., TAKEUCHI T.;
RT Biochemical and molecular characterization of nucleoside
RT triphosphate hydrolase isozymes from the parasitic protozoan
RT Toxoplasma gondii."
RL J. Biol. Chem. 270:11391-11397(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-RH;
RX MEDLINE; 95050750.
RA BERAMDES D., PECK K.R., AFIFI M.A., BECKERS C.J.M., JOINER K.A.;
RT "Randomly repeated genes encode nucleoside triphosphate hydrolase
RT isoforms secreted into the parasitophorous vacuole of Toxoplasma
RT gondii."
RL J. Biol. Chem. 269:29252-29260(1994).
RN [3]
RP SEQUENCE OF 334-523 FROM N.A.
RC STRAIN-RH;
RX MEDLINE; 90152366.
RA JOHNSON A.M., ILLANA S., MCDONALD P.J., ASAI T.;
RT "Cloning, expression and nucleotide sequence of the gene fragment
RT encoding an antigenic portion of the nucleoside triphosphate
RT hydrolase of Toxoplasma gondii."
RL Gene 85:215-220(1989).
CC -1- FUNCTION: MAY PERFORM AN IMPORTANT PROCESSING STEP IN THE
CC CONVERSION OF HIGH ENERGY NUCLEOTIDES PRIOR TO UPTAKE BY THE
CC PARASITE. NTPASE-II HAS A SPECIFIC ACTIVITY 4.5-FOLD LOWER THAN

```

```

CC NTPASE-I IN HYDROLYSIS OF ATP. THE PRIMARY DIFFERENCE BETWEEN
CC THESE ISOZYMES LIES IN THEIR ABILITY TO HYDROLYZE NUCLEOSIDE
CC TRIPHOSPHATE VERSUS DIPHOSPHATE SUBSTRATES. WHILE NTPASE-II
CC HYDROLYZES ADP TO ADP AND ADP TO AMP AT ALMOST THE SAME RATE,
CC NTPASE-I HYDROLYZES ADP TO AMP AT A MUCH SLOWER RATE (0.7% OF THE
CC RATE FOR ATP).
CC -1- CATALYTIC ACTIVITY: NTP + H(2)O = NDP + PHOSPHATE.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: SECRETED; FOUND IN HOST CELL PARASITOPHOUS
CC VACUOLE.
CC -1- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L39077; AAC41569.1; -.
CC EMBL; L39079; AAC41570.1; -.
CC EMBL; U96965; AAC80187.1; -.
CC EMBL; M33472; AAC30143.1; -.
CC PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
CC DR PFAM; PF01150; GDA1_CD39; 1.
CC KW Hydrolase; Multigene family; Signal.
CC FT SIGNAL 1 25
CC FT CHAIN 26 628
CC FT CARBOHYD 432 432
CC FT VARIANT 6 628
CC FT VARIANT 91 91
CC FT VARIANT 101 101
CC FT VARIANT 334 334
CC FT CONFLICT 370 370
CC FT CONFLICT 372 372
CC FT CONFLICT 399 399
CC FT CONFLICT 437 437
CC FT CONFLICT 488 489
CC FT CONFLICT 492 494
CC FT CONFLICT 497 497
CC FT CONFLICT 499 499
CC FT CONFLICT 523 523
CC FT CONFLICT 523 523
CC SEQUENCE 628 AA; 69585 MW; DABA665D CRC32;

Query Match 3.9%; Score 127; DB 1; Length 628;
Best Local Similarity 32.8%; Pred. No. 4,46e-05;
Matches 22; Conservative 21; Mismatches 22; Indels 2; Gaps 2;

DB 228 TRPIGAEGLFAFTTLNHLRSRLGEDPARCMIDEGVGHCRNDLAGVVEGASAOIVF 287
QY 129 ARITGOERGAGWITTYLLGKFSQKTRMFSIYEPYETNN-QETP-CALDGGASTQVTF 186
DB 288 PLOGCTV 294
QY 187 VPQNOTI 193

RESULT 10
ID NTP4_TOXGO STANDARD; PRT; 592 AA.
AC P52913;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1999 (Rel. 39, Last annotation update)
DE PUTATIVE NUCLEOSIDE-TRIPHOSPHATASE (EC 3.6.1.15) (NUCLEOSIDE
DE TRIPHOSPHATE PHOSPHOHYDROLASE) (NTPASE).
GN NTP4.
OS Toxoplasma gondii.
CC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Sarcocystidae;
CC Toxoplasma.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NICOLE;

```

RA COTE S., MORENCY M., LEVESQUE R.C.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: NTP + H(2)O = NDP + PHOSPHATE.
CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.
CC -1- CAUTION: THIS COULD BE A PSEUDOGENE. IT DOES NOT SEEM TO START
WITH AN INITIATOR MET.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U28353; AAA80336.1; ALT_INIT.
DR PROSITE; PS01238; GDAL_CD39_NTPASE; 1.
DR PFAM; PF01150; GDAL_CD39; 1.
KW Hydrolase; Multigene family.
SQ SEQUENCE 592 AA; 65762 MW; DDD65858 CRC32;

Query Match 3.8%; Score 125; DB 1; Length 592;
Best Local Similarity 32.8%; Pred. No. 9,44e-05;
Matches 22; Conservative 20; Mismatches 23; Indels 2; Gaps 2;

DB 192 TRPTGAEGLFAFTLTHLSRLGDEDPARCMIDERYGKOCRNLAGVVEGASAOIVF 251
QY 129 ARIITGEGAGWITTYLLGKFSQKTRWFSIVYETNN-QEYF-GALDYGASTOYTF 186
DB 252 PLOEGTV 258
QY 187 VPQNGTI 193

RESULT 11
ID NTP1_TOXGO STANDARD; PRT; 628 AA.
AC 027893;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE NUCLEOSIDE-TRIPHOSPHATASE I PRECURSOR (EC 3.6.1.15) (NTPASE-I)
DE (NUCLEOSIDE TRIPHOSPHATE HYDROLASE 1).
GN NTP3.
OS Toxoplasma gondii.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Sarcocystidae;
OC Toxoplasma.
OC [1]
RN SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.
RP STRAIN-RH;
RC MEDLINE; 95263459.
RA ASAI T., MIDRA S., SIBLEY L.D., OKABAYASHI H., TAKEUCHI T.;
RT "Biochemical and molecular characterization of nucleoside
RT triphosphate hydrolase isozymes from the parasitic protozoan
RT Toxoplasma gondii.";
RL J. Biol. Chem. 270:11391-11397(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-RH;
RC MEDLINE; 95050750.
RA BERNHEDS D., PECK K.R., AFIPI M.A., BECKERS C.J.M., JOINER K.A.;
RT "Randomly repeated genes encode nucleoside triphosphate hydrolase
RT isoforms secreted into the parasitophorous vacuole of Toxoplasma
RT gondii.";
RL J. Biol. Chem. 269:29252-29260(1994).
CC -1- FUNCTION: MAY PERFORM AN IMPORTANT PROCESSING STEP IN THE
CC CONVERSION OF HIGH ENERGY NUCLEOTIDES PRIOR TO UPTAKE BY THE
CC PARASITE AND MAY CONTRIBUTE TO INTRACELLULAR SURVIVAL AND
CC VIRULENCE. NTPASE-I HAS A SPECIFIC ACTIVITY 4.5-FOLD HIGHER THAN
CC NTPASE-II IN HYDROLYSIS OF ATP. THE PRIMARY DIFFERENCE BETWEEN
CC THESE ISOZYMES LIES IN THEIR ABILITY TO HYDROLYZE NUCLEOSIDE
CC TRIPHOSPHATE VERSUS DIPHOSPHATE SUBSTRATES. WHILE NTPASE-II
CC HYDROLYZES ATP TO ADP AND AMP AT ALMOST THE SAME RATE,
CC NTPASE-I HYDROLYZES ADP TO AMP AT A MUCH SLOWER RATE (0.7% OF THE

CC RATE FOR ATP).
CC -1- CATALYTIC ACTIVITY: NTP + H(2)O = NDP + PHOSPHATE.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: SECRETED; FOUND IN HOST CELL PARASITOPHOUS
CC VACUOLE.
CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L39078; AAA89203.1; -;
DR EMBL; U96965; AAC80188.1; -;
DR PROSITE; PS01238; GDAL_CD39_NTPASE; 1.
DR PFAM; PF01150; GDAL_CD39; 1.
KW Hydrolase; Multigene family; Signal.
FT SIGNAL 1 25
FT CHAIN 26 628 NUCLEOSIDE-TRIPHOSPHATASE I.
FT CARBOHYD 432 432 POTENTIAL.
SQ SEQUENCE 628 AA; 69159 MW; DA2A1577 CRC32;

Query Match 3.8%; Score 125; DB 1; Length 628;
Best Local Similarity 32.8%; Pred. No. 9,44e-05;
Matches 22; Conservative 20; Mismatches 23; Indels 2; Gaps 2;

DB 228 TRPTGAEGLFAFTLTHLSRLGDEDPARCMIDERYGKOCRNLAGVVEGASAOIVF 287
QY 129 ARIITGEGAGWITTYLLGKFSQKTRWFSIVYETNN-QEYF-GALDYGASTOYTF 186
DB 288 PLOEGTV 294
QY 187 VPQNGTI 193

RESULT 12
ID UDRA_HUMAN STANDARD; PRT; 528 AA.
AC P36537;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE UDP-GLUCONOSYLTTRANSFERASE 2B10 PRECURSOR, MICROSOMAL (EC 2.4.1.17)
DE (UDPGR).
GN UGT2B10.
GN HGT2B10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Euteria; Primates; Catarrhini; Hominiidae; Homo.
OC [1]
RN SEQUENCE FROM N.A.
RP TISSUE-LIVER;
RC MEDLINE; 93326164.
RA JIN C.-J., MINERS J.O., LILLYWHITE K.J., MACKENZIE P.I.;
RT "cDNA cloning and expression of two new members of the human liver
RT UDP-glucuronosyltransferase 2B subfamily.";
RL Biochem. Biophys. Res. Commun. 194:496-503(1993).
CC -1- FUNCTION: UDPGR IS OF MAJOR IMPORTANCE IN THE CONJUGATION AND
CC SUBSEQUENT ELIMINATION OF POTENTIALLY TOXIC XENOBIOTICS AND
CC ENDOGENOUS COMPOUNDS.
CC -1- CATALYTIC ACTIVITY: UDP-GLUCURONATE + ACCEPTOR = UDP + ACCEPTOR
CC BETA-D-GLUCURONOSIDE.
CC -1- SUBCELLULAR LOCATION: MICROSOMAL.
CC -1- SIMILARITY: BELONGS TO THE UDP-GLYCOSYLTTRANSFERASE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

```

CC -----
DR EMBL; X63359; CAA44961.1; -
DR PIR; JN0620; JN0620.
DR MIM; 600070; -
DR PROSITE; PS00375; UDPGT; 1.
DR PFAM; PF00201; UDPGT; 1.
DR Transferrase; Glycosyltransferase; Glycoprotein; Transmembrane; Signal;
KW Multigene family; Microsome.
FT SIGNAL 1 23 BY SIMILARITY.
FT CHAIN 1 23 UDP-GLUCURONOSYLTRANSFERASE 2B10.
FT TRANSFER 492 512 POTENTIAL.
FT CARBOHYD 66 66 POTENTIAL.
FT CARBOHYD 314 314 POTENTIAL.
FT CARBOHYD 481 481 POTENTIAL.
SQ SEQUENCE 528 AA; 60774 MW; 63F280A6 CRC32;

Query Match 3.6%; Score 117; DB 1; Length 528;
Best Local Similarity 24.7%; Pred. No. 1.77e-03;
Matches 21; Conservative 24; Mismatches 36; Indels 4; Gaps 4;

Db 133 KKLMLKLOSREDIYVADAYLPCGELLAELEFNPV-YSH-SFSPGYSFRRHSGGF-IFP 189
QY 266 KHEFTLPPQOFELQIGVYQCHOSILELFTNTSYCPYSOCAPNGIF-LPPLGDFGAFS 324
ID TBA3_MAIZE STANDARD; PRT; 450 AA.
AC P22275;
DT 01-AUG-1991 (Rel. 19; Created)
DT 01-AUG-1991 (Rel. 19; Last sequence update)
DT 15-JUL-1999 (Rel. 38; Last annotation update)
DE TUBULIN ALPHA-3 CHAIN.
GN TUBA3 OR TUA3.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
OC Poaceae; Zea.
CC [1]
RN SEQUENCE FROM N.A.
RP MEDLINE; 91078640.
RA MONTOLIU L., PUIGDOMENECH P., RIGAU J.;
RT "The Tuba alpha 3 gene from Zea mays: structure and expression in
RT dividing plant tissues.";
RL Gene 94:201-207(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. B73; TISSUE=SHOOT;
RX MEDLINE; 92395680.
RA VILLEWOR R., JOYE C.M., HAAS N.A., GODDARD R.H., KOPCEK S.D.,
RA HUSSEY P.J., SNUSTAD D.P., SIEFLOW C.D.;
RT "Alpha-tubulin gene family of maize (Zea mays L.). Evidence for two
RT ancient alpha-tubulin genes in plants.";
RL J. Mol. Biol. 227:81-96(1992).
CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL; M60171; AAA33518.1; -
DR EMBL; X63176; CAA44861.1; -

```

```

DR PIR; JN0105; JN0105.
DR PIR; S28979; S28979.
DR MAIZEDB; 17141; -
DR MENDEL; 1410; ZEAMA:TUBA.3.
DR PROSITE; PS00227; TUBULIN; 1.
DR PFAM; PF00091; tubulin; 1.
KW Microtubules; GTP-binding; Multigene family.
FT NP BIND 142 148 GTP (POTENTIAL).
FT SITE 450 450 INVOLVED IN POLYMERIZATION.
SQ SEQUENCE 450 AA; 49561 MW; 6E7B5922 CRC32;

Query Match 3.1%; Score 102; DB 1; Length 450;
Best Local Similarity 26.4%; Pred. No. 2.99e-01;
Matches 19; Conservative 24; Mismatches 25; Indels 4; Gaps 4;

Db 43 GHDDAFSTFSGTGAKVPAIFVDEPYIDEV-RT-GYRQLFHPQOLISGEDAA 100
QY 82 SOHET-PVYLQAT-AGMLRLESEELADVLVVERSLSNYPDFGQARLTGQEGCA 139
ID SVH_ORYZA STANDARD; PRT; 494 AA.
AC P93422;
DT 15-DEC-1998 (Rel. 37; Created)
DT 15-DEC-1998 (Rel. 37; Last sequence update)
DT 15-DEC-1998 (Rel. 37; Last annotation update)
DE HISTIDYL-TRNA SYNTHETASE (EC 6.1.1.21) (HISTIDINE--TRNA LIGASE)
DE (HISRS).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
OC Poaceae; Oryza.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA AKASHI K., SMALL I.D.;
RT "A cDNA clone encoding rice histidyl-trna synthetase.";
RL (in) Plant Gene Register PGR97-062.
CC -1- CATALYTIC ACTIVITY: ATP + L-HISTIDINE + TRNA(HIS) = AMP +
CC PYROPHOSPHATE + L-HISTIDYL-TRNA(HIS).
CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL; Z85984; CAB06553.1; -
DR PROSITE; PS00179; AA_TRNA_LIGASE_II.1; 1.
DR PROSITE; PS00339; AA_TRNA_LIGASE_II.2; 1.
DR PFAM; PF00587; tRNA-synt_2b; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
SQ SEQUENCE 494 AA; 55326 MW; 98E27043 CRC32;

Query Match 3.1%; Score 102; DB 1; Length 494;
Best Local Similarity 19.8%; Pred. No. 2.99e-01;
Matches 23; Conservative 42; Mismatches 40; Indels 11; Gaps 9;

Db 128 CSLRYDLYVFPARYAMNSSLKRYQAKVYRRNPSPKGRYREFYQDFDIAGYV-ETM 186
QY 171 KHEFTLPPQOF-ELQIGVYQCHOS-IL--E-LFTNTSYCPYSQCAFN--GIFLPL 316
ID EPDEVIKAVTELDDLDIGTEYIKLNHRKLLDGLMLICGVPEKFRVYCSGI-DK 241
QY 317 QGDFGAFSAFYVMKFLNLT-S-E-KVSEKVTENMKKCAOPWMEIKRYSVAGVKK 370

```


RESULT 15
ID UDPF_HUMAN STANDARD; PRT: 530 AA.
AC P54855; P23765;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE UDP-GLUCURONOSYLTRANSFERASE 2B15 PRECURSOR, MICROSOMAL (EC 2.4.1.17)
DE (UDPGT) (UDPETH-3) (HUG44).
GN UGT2B15 OR UGT2B8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LIVER;
RA MEDLINE; 95136867.
RA GREEN M.D., OTURU E.M., TEPHLY T.R.;
RT "Stable expression of a human liver UDP-glucuronosyltransferase
(UGT2B15) with activity toward steroid and xenobiotic substrates.";
RL Drug Metab. Dispos. 22:799-805(1994).
RN [2]
RP SEQUENCE OF 8-530 FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE-LIVER;
RA MEDLINE; 90343358.
RA COFFMAN B.L., TEPHLY T.R., IRSHAD Y.M., GREEN M.D., SMITH C.,
RA JACKSON M.R., WOOSTER R., BURCHELL B.;
RT "Characterization and primary sequence of a human hepatic microsomal
estriol UDPglucuronosyltransferase.";
RL Arch. Biochem. Biophys. 281:170-175(1990).
RN [3]
RP VARIANT TYR-85.
RA MEDLINE; 97439504.
RA LEVESQUE E., BEAULIEU M., GREEN M.D., TEPHLY T.R., BELANGER A.,
RA HDW D.W.;
RT "Isolation and characterization of UGT2B15(Y85): a
UDP-glucuronosyltransferase encoded by a polymorphic gene.";
RL Pharmacogenetics 7:317-325(1997).
CC -1- FUNCTION: UDPGT IS OF MAJOR IMPORTANCE IN THE CONJUGATION AND
SUBSEQUENT ELIMINATION OF POTENTIALLY TOXIC XENOBIOTICS AND
ENDOGENOUS COMPOUNDS. THIS ISOZYME DISPLAYS ACTIVITY TOWARD
SEVERAL CLASSES OF XENOBIOTIC SUBSTRATES, INCLUDING SIMPLE
PHENOLIC COMPOUNDS, 7-HYDROXYLATED COCAINE, FLAVONOIDS,
ANTHRAQUINONES, AND CERTAIN DRUGS AND THEIR HYDROXYLATED
METABOLITES. IT ALSO CATALYZES THE GLUCURONIDATION OF ENDOGENOUS
ESTROGENS AND ANDROGENS.
CC -1- CATALYTIC ACTIVITY: UDP-GLUCURONATE + ACCEPTOR = UDP + ACCEPTOR
BETA-D-GLUCORONOSIDE.
CC -1- SUBCELLULAR LOCATION: MICROSOMAL.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN MANY TISSUES.
CC -1- SIMILARITY: BELONGS TO THE UDP-GLYCOSYLTRANSFERASE FAMILY.
CC -1- CAUTION: REF.2 THOUGHT THAT THIS WAS A SEPARATE FORM (UGT2B8). THE
NAME UGT2B8 HAS NOW BEEN REUSED FOR A RAT ENZYME.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U08854; AAC50077.1; -;
DR EMBL; U06641; AAA83406.1; -;
DR PIR; S11309; S11309.
DR MIM; 600069; -;
DR PROSITE; PS00375; UDPGT; 1.
DR PRAM; PF00201; UDPGT; 1.
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane; Signal;
KW Multigene family; Microsome; Polymorphism.
FT SIGNAL 1 23
FT CHAIN 24 530 UDP-GLUCURONOSYLTRANSFERASE 2B15.
FT TRANSMEM 495 515 POTENTIAL.
FT

FT CARBOHYD 65 65 POTENTIAL.
FT CARBOHYD 316 316 POTENTIAL.
FT CARBOHYD 483 483 POTENTIAL.
FT VARIANT 85 85
FT
FT CONFLICT 119 119 /FtId=VAR.007713.
FT CONFLICT 145 145 E -> A (IN REF. 2).
FT CONFLICT 150 155 K -> R (IN REF. 2).
FT CONFLICT 162 162 LADALN -> PGDFVF (IN REF. 2).
FT CONFLICT 165 165 A -> S (IN REF. 2).
FT CONFLICT 170 177 F -> L (IN REF. 2).
FT CONFLICT 181 181 LYSLEFSV -> YRSRSR (IN REF. 2).
FT CONFLICT 203 203 S -> I (IN REF. 2).
FT CONFLICT 293 293 M -> L (IN REF. 2).
FT CONFLICT 401 401 H -> D (IN REF. 2).
FT CONFLICT 443 443 V -> A (IN REF. 2).
FT CONFLICT 501 501 C -> W (IN REF. 2).
FT CONFLICT 523 523 T -> K (IN REF. 2).
SQ SEQUENCE 530 AA; 60961 MM; 78E9198E CRC32;

Query Match 3.1%; Score 102; DB 1; Length 530;
Best Local Similarity 21.9%; Pred. No. 2.99e-01;
Matches 21; Conservative 27; Mismatches 44; Indels 4; Gaps 4;

Db 135 KIKMKLOESKFDVYLADALNPGCELLAELEFNIPF-LTSL-RESYGYFEKNKGGE-LFP 191
QY 266 KRFETLPPQPFETIOGIGNYOQCHOSTLELFTSTCPYSQCAFNGIF-LPPLQDGFAGFS 324
Db 192 PSYPPVWSESDQMIEMERIKMIMHLYPFWFQI 227
QY 325 AFYFMKFLNLTSEKVSQEKTEMMKRFCAQPMERI 360

Search completed: Fri May 5 09:26:47 2000
Job time : 174 secs.


```

FT DOMAIN 38 478 EXTRACELLULAR (POTENTIAL).
FT TRANSEM 479 499 POTENTIAL.
FT DOMAIN 500 511 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 73 73 POTENTIAL.
FT CARBOHYD 226 226 POTENTIAL.
FT CARBOHYD 291 291 POTENTIAL.
FT CARBOHYD 333 333 POTENTIAL.
FT CARBOHYD 374 374 POTENTIAL.
FT CARBOHYD 429 429 POTENTIAL.
FT CARBOHYD 458 458 POTENTIAL.
SQ SEQUENCE 511 AA; 57408 MW; E59D8C9 CRC32;

Query Match 80.1%; Score 2624; DB 11; Length 511;
Best Local Similarity 74.4%; Pred. No. 0.00e+00;
Matches 328; Conservative 63; Mismatches 47; Indels 3; Gaps 3;

Db 38 THNKLPEVNVKGYILDAGSSHTSLYIKWPAKENDTGVTQVECECRVKGPGISKFYOK 97
1 TONKALPENVKGYILDAGSSHTSLYIKWPAKENDTGVTQVECECRVKGPGISKFYOK 60

Db 98 TDETAAYLAECMKMSTERIPASKOHPYLGATAGMRLRMESKOSADEVLAASRSIK 157
61 VNEIGIYLTDCERAREVYIPRSOHOETPYLGATAGMRLRMESSELDADYDVERSL 120

Db 158 SYPDFOGAKITGOEGAGYGMTINYLGRFTOESWMLNFTS-DSOKATEGALDLGGS 216
121 NYPDFOGAKRITGOEGAGYGMTINYLGRFSQKTRMFSIYPTENNQETFGALDLGGA 180

Db 217 STQTFVPLNLTLEAPESLQRLKGYTYVTHSFLCYGKQALMOKLADIOVSSGI 276
181 STQTFVPLNLTLEAPESLQRLKGYTYVTHSFLCYGKQALMOKLADIOVSSGI 240

Db 277 LKDCPEFYKRVVNSVSELYGPTCKREFEKLRFNOFQVQGTGDEQCHQSILKFPNNH 336
241 LRDCPEFYKRVVNSVSELYGPTCKREFEKLRFNOFQVQGTGDEQCHQSILKFPNNH 300

Db 337 CPYSQCAFNGVFLPPLGSGFSAFYVMDFFKKMANSVSSQEKMEITNFCSPWE 396
301 CPYSQCAFNGVFLPPLGSGFSAFYVMDFFKKMANSVSSQEKMEITNFCSPWE 358

Db 397 EVKASYPVKKRYLSEYCFSGTYILSLQGNFTGTSMDQHFPMKIDSNAGWTLG 456
359 EIKTSYAVGKRYLSEYCFSGTYILSLQGNFTGTSMDQHFPMKIDSNAGWTLG 418

Db 457 LNTNMIAPBQPLSPPLHST 477
419 LNTNMIAPBQPLSPPLHST 439

RESULT 2 PRELIMINARY; PRT; 513 AA.
AC 018956;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
DE VASCULAR ATP-DIPHOSPHODROLASE (EC 3.6.1.5) (LYMPHOID CELL ACTIVATION
DE ANTIGEN) (CD39 ANTIGEN) (ATPASE) (ECTO-ATPASE) (ECTO-
DE ADPASE) (NUCLEOTIDE PHOSPHODROLASE) (ATP PYROPHOSPHODROLASE).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
RN [1]
RP TISSUE-AORTIC ENDOTHELIUM.
RA CHANG A.S., GARCIA R.L., CHANG S.M., SCHILLING W.P.;
RL Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: HYDROLASES EXTRACELLULAR ATP AND ADP TO AMP. COULD
CC INHIBIT PLATELET AGGREGATION IN RESPONSE TO ADP, COLLAGEN, AND
CC THROMBIN (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 PHOSPHATE.
CC -1- COFACTOR: CALCIUM ION-AND MAGNESIUM ION-DEPENDENT ACTIVITY (BY
CC SIMILARITY).

```

```

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
DR EMBL: AF005940; AAB62382.1; -.
DR PROSITE: PS01238; GDA1_CD39_NTPASE; 1.
DR PFAM: PF01150; GDA1_CD39; 1.
KW Hydroxylase; Transmembrane; Antigen; Glycoprotein.
FT DOMAIN 1 16 CYTOPLASMIC (POTENTIAL).
FT TRANSEM 17 37 POTENTIAL.
FT DOMAIN 38 481 EXTRACELLULAR (POTENTIAL).
FT TRANSEM 482 502 POTENTIAL.
FT DOMAIN 503 513 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 73 73 POTENTIAL.
FT CARBOHYD 227 227 POTENTIAL.
FT CARBOHYD 245 245 POTENTIAL.
FT CARBOHYD 307 307 POTENTIAL.
FT CARBOHYD 336 336 POTENTIAL.
FT CARBOHYD 373 373 POTENTIAL.
FT CARBOHYD 460 460 POTENTIAL.
SQ SEQUENCE 513 AA; 58113 MW; 17735933 CRC32;

Query Match 72.9%; Score 2389; DB 6; Length 513;
Best Local Similarity 69.8%; Pred. No. 0.00e+00;
Matches 309; Conservative 71; Mismatches 58; Indels 5; Gaps 4;

Db 38 TONKALPENVKGYILDAGSSHTSLYIKWPAKENDTGVTQVEESNVKGPISGFPAK 97
1 TONKALPENVKGYILDAGSSHTSLYIKWPAKENDTGVTQVEESNVKGPISGFPAK 60

Db 98 VNEINYLTLACERAKVYIPSIQHMETPYLGATAGMRLRMENKQAMDKILAAVASSIS 157
61 VNEIGIYLTDCERAREVYIPRSOHOETPYLGATAGMRLRMESSELDADYDVERSL 120

Db 158 EYPPDFOGAKRITGOEGAGYGMTINYLGRFTOESWMLNFTS-DSOKATEGALDLGGA 217
121 NYPDFOGAKRITGOEGAGYGMTINYLGRFSQKTRMFSIYPTENNQETFGALDLGGA 180

Db 217 STQTFVPLNLTLEAPESLQRLKGYTYVTHSFLCYGKQALMOKLADIOVSSGI 276
181 STQTFVPLNLTLEAPESLQRLKGYTYVTHSFLCYGKQALMOKLADIOVSSGI 240

Db 277 LKDCPEFYKRVVNSVSELYGPTCKREFEKLRFNOFQVQGTGDEQCHQSILKFPNNH 336
241 LRDCPEFYKRVVNSVSELYGPTCKREFEKLRFNOFQVQGTGDEQCHQSILKFPNNH 300

Db 337 CPYSQCAFNGVFLPPLGSGFSAFYVMDFFKKMANSVSSQEKMEITNFCSPWE 396
301 CPYSQCAFNGVFLPPLGSGFSAFYVMDFFKKMANSVSSQEKMEITNFCSPWE 358

Db 397 EVKASYPVKKRYLSEYCFSGTYILSLQGNFTGTSMDQHFPMKIDSNAGWTLG 456
359 EIKTSYAVGKRYLSEYCFSGTYILSLQGNFTGTSMDQHFPMKIDSNAGWTLG 418

Db 457 LNTNMIAPBQPLSPPLHST 477
419 LNTNMIAPBQPLSPPLHST 439

RESULT 3 PRELIMINARY; PRT; 306 AA.
AC 09Y309;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
DE ECTO-ATP DIPHOSPHODROLASE II (EC 3.6.1.5).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RP TISSUE-PLACENTA;
RA MATSUMOTO M., SAKURAI Y., KOKUBO T., YAGI H., MATSUI T., TITANI K.,
RA FUJIMURA Y., NARITA N.;

```

"The cDNA cloning of human placental ecto-ATP diphosphohydrolases I and II.";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ133134; CAB41867.1; -
 DR PROSITE; PS01238; GDAL_CD39_NTPASE; 1.

KW Hydrolase.
 SO SEQUENCE 306 AA; 34175 MW; 2F3C431 CRC32;

Query Match 52.3%; Score 1712; DB 4; Length 306;
 Best Local Similarity 95.9%; Pred. NO. 0.00e+00;

Matches 236; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

Db 45 TONALPENNYGIVLADGSSHTSLYIKWPAEKENDGVVHOVEECVKGPGISKFYOK 104
 |||||
 QY 1 TONALPENNYGIVLADGSSHTSLYIKWPAEKENDGVVHOVEECVKGPGISKFYOK 60
 |||||
 Db 105 VNEIGIYLTDCMERAREVYIPRSOHOETPVYLGATAGMRLMESEBLADRYLDVVERSL 164
 |||||
 QY 61 VNEIGIYLTDCMERAREVYIPRSOHOETPVYLGATAGMRLMESEBLADRYLDVVERSL 120
 |||||
 Db 165 NYPPDFOGARITITGOEGAGYMTITNLLGFSQKTRMFSTVPTETNNQETFGALDLGGA 224
 |||||
 QY 121 NYPPDFOGARITITGOEGAGYMTITNLLGFSQKTRMFSTVPTETNNQETFGALDLGGA 180
 |||||
 Db 225 STQVTFVPOONTIESPDNALQFRLYGKDYNYTTHSFLCYGKDALMOKLADIDIASITQS 284
 |||||
 QY 181 STQVTFVPOONTIESPDNALQFRLYGKDYNYTTHSFLCYGKDALMOKLADIDIASITQS 240
 |||||
 Db 285 -RPAPF 289
 |||||
 QY 241 LRDPGF 246

RESULT 4
 ID 093295 PRELIMINARY; PRT; 493 AA.

AC 093295;
 DT 01-NOV-1998 (TREMblrel. 08, Created)
 DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
 DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)
 DE ECTO-ATP-DIPHOSPHOHYDROLASE.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
 OC Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=OVIDUCT;
 RX MEDLINE; 98298108.
 RA NAGY A.K., KNOWLES A.F., NAGAMI G.T.;
 RT "Molecular cloning of the chicken oviduct ecto-ATP-
 diphosphohydrolase.";
 RL J. Biol. Chem. 273:16043-16049(1998).
 DR EMBL; AF041355; AAC26491.1; -
 DR PROSITE; PS01238; GDAL_CD39_NTPASE; 1.
 DR PFM; PFO1150; GDAL_CD39; 1.
 KW Hydrolase.
 SO SEQUENCE 493 AA; 54034 MW; A6E48E68 CRC32;

Query Match 44.4%; Score 1453; DB 13; Length 493;
 Best Local Similarity 45.5%; Pred. NO. 0.00e+00;

Matches 194; Conservative 88; Mismatches 134; Indels 10; Gaps 10;

Db 35 LPBGTGKGLVDASTHTALYVOMPADKENGITGVSQVESCCTYNGSGTSSYADDPAGAG 94
 |||||
 QY 6 LPBWTGKGLVDASTHTALYVOMPADKENGITGVSQVESCCTYNGSGTSSYADDPAGAG 65
 |||||
 Db 95 ASLRPCDKAAVAVPEOQOTPTLYGTATAGMRLREONSKAOVFAEVSKAIREPPVD 154
 |||||
 QY 66 IYLDCHERAREVYIPRSOHOETPVYLGATAGMRLMESEBLADRYLDVVERSLSNPFD 125
 |||||
 Db 155 FRGAQILTNGEGSGFTVYLLLETLI-KFS-PA-GKMEHPONTTEGALDLGASTQI 211
 |||||
 QY 126 FQGARITITGOEGAGYMTITNLLGFSQKTRMFSTVPTETNNQETFGALDLGASTQI 184
 |||||

Db 212 TFPQ-GLTIEDKNMSTVLEFLYGTNGSLYTHSLCYGQIOA-SKRMLAALHODGSYVONIS 269
 |||||
 QY 185 TFPQONTIESPDNALQFRLYGKDYNYTTHSFLCYGKDALMOKLAKOI-OVASN-ETLR 242
 |||||
 Db 270 HPCYPKGYRRIITTAETIDSCVPTPSMISPAQILTYGTGNPACPAIILKFLNLTGCA 329
 |||||
 QY 243 DPCFHPGKRVANNVSDLYKTPCTKRFEMTLFPQGEIIGICIGNYQCHOSILELFTSYCP 302
 |||||
 Db 330 NRTGCFDGVYPPVARGOFPFAFAGFYTFSTNLNGOO-SLSHVATVDFCNKMSLVE 388
 |||||
 QY 303 YSOCAFNGIFLPLDGDGFAFAFYVAKFLNLSEKVSQEVTEEMKRFCAQPFEEIKT 362
 |||||
 Db 389 TFPQNK-GLTHTCYVGVYIITLLVDGKFKDEHMSNTHFQKAGNADIGTGLNLT 447
 |||||
 QY 363 SYAGVKEKYLSECHPSGYIISLLDGYHFTADSWERHIFGKIQGSDAGWTLGYMLNT 422
 |||||
 Db 448 NMIPTE 453
 |||||
 QY 423 NMIPAE 428

RESULT 5
 ID 055026 PRELIMINARY; PRT; 495 AA.

AC 055026; 035928;
 DT 01-JUN-1998 (TREMblrel. 06, Created)
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
 DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)
 DE ECTO-ATPASE (EC 3.6.1.5).
 GN CD39L1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE; 98288263.
 RA GAO L., DONG L., WHITLOCK J.P., JR.;
 RT "A novel response to dioxin. Induction of ecto-ATPase gene
 expression.";
 RL J. Biol. Chem. 273:15358-15365(1998).
 RN [2]
 RP SEQUENCE FROM N.A. (SHORT FORM).
 RC TISSUE=EMBRYO;
 RX MEDLINE; 97419269.
 RA CHADWICK B.P., FRISCHAUF A.M.;
 RT "Cloning and mapping of a human and mouse gene with homology to ecto-
 ATPase genes.";
 RL Mamm. Genome 8:668-672(1997).

CC -1- FUNCTION: HYDROLYSES EXTRACELLULAR ATP TO AMP.
 CC -1- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 PHOSPHATE.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- ALTERNATIVE PRODUCTS: TWO ISOFORMS; A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- INDUCTION: BY DIOXIN.
 CC -1- PTM: HAS PROBABLY VARIOUS DISULFIDE-BONDS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.

DR EMBL; AF042811; AAC24347.1; -
 DR EMBL; U91511; AAB81014.1; -
 DR MGD; MGI:1096863; C63911.
 DR PROSITE; PS01238; GDAL_CD39_NTPASE; 1.
 DR PFM; PFO1150; GDAL_CD39; 1.
 KW Hydrolase; Transmembrane; Antigen; Glycoprotein; Alternative splicing.
 SO CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 1 4
 FT TRANSMEM 1 25
 FT DOMAIN 22 462
 FT TRANSMEM 26 462
 FT DOMAIN 463 483
 FT TRANSMEM 467 470
 FT DOMAIN 484 495
 FT VARSPIC 130 132
 FT VARSPIC 133 495
 FT CARBOHYD 64 64
 FT CARBOHYD 129 129

FT POLY-LEU. EXTRACELLULAR (POTENTIAL).
 FT POLY-LEU. CYTOPLASMIC (POTENTIAL).
 FT LYS -> MAG (IN SHORT FORM).
 FT MISSING (IN SHORT FORM).
 FT POTENTIAL.
 FT POTENTIAL.

FT CARBOHYD 294 294 POTENTIAL.
 FT CARBOHYD 319 319 POTENTIAL.
 FT CARBOHYD 378 378 POTENTIAL.
 FT CARBOHYD 443 443 POTENTIAL.
 SQ SEQUENCE 495 AA; 54310 MW; B9CE5702 CRC32;
 Query Match 41.8%; Score 1368; DB 11; Length 495;
 Best Local Similarity 45.4%; Pred. No. 9.17e-286;
 Matches 198; Conservative 86; Mismatches 135; Indels 17; Gaps 15;
 Db 29 TODVREPPALKGIYVDAGSSHTSMFYKMPADKENDTGIVGSHSCDVGAGGSISSAND 88
 QY 1 TONKALPENVKKGIYVDAGSSHTSLYIKMPAKENDGVHVECECKYKGGISKVYOK 60
 Db 89 PSRAGOSLVCELEOALRDY-PKDRVASTPLYLGATAGMRPLNTSPKATVLEAVTOTL 147
 QY 61 VNEIGIYILDCEMERA-REVIPIRSQHETPVYLGATAGMRLNMBESBELADRYLDYVERSL 119
 Db 148 TRYPPDFRGARILISGODEVFGWNTANYLLENFI-KYGV-VGRWIRPKRGLGAMDLCG 204
 QY 120 SNYPDFOGARILITGEEGAYGMITINYLKFKSQKTRMFSIYVETNNQETFGALDLCG 179
 Db 205 ASTQTFETTS-EDPQNEVHLRLYGQHYRYTHSFLCYGQDQVL-QRLASALDI--H 260
 QY 180 ASTQVTFPQONQITSPDUALQFLRGKDYNYTHSFLCYGQDQALMOKL-ARDIQVANS 238
 Db 261 RF-H-PCMPKGYSTOVLREYVOSPTMGORPOTFNSAIVSLSGTSMALCDIVSGLF 318
 QY 239 ELIRPCHFGKRYKVVNSDLTKTPCTK-REFEMLT-PFOOFELQIGNTQOCHSILELF 296
 Db 319 NISSCFSCSPGNGVFOFPVAGNFIAFSAFYTVDFLTVMGLPVTCLKOLEBATEITCN 378
 QY 297 NTSYCPYSCAFNGIFLPLQDGFAGFSAFYFMKFLN-LTSEKVSQEK-VTEMKMKFCA 354
 Db 379 QTMELQARVPQOQTR-LPDYCTVAMFIHQLSRGYSFEDRSRGRVYFEKKAADYAVGMT 437
 QY 355 QWBEIKTSYAGVKEKYLSEYCFSGTYILSLQGYHFTADSWEHITHFGIKIGSDAGMTL 414
 Db 438 LGYMLNTNLIPADLP 453
 QY 415 LGYMLNTNLIPADLP 430
 RESULT 6 PRELIMINARY; PRT; 495 AA.
 ID O35795;
 AC O35795;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
 DE ECTO-ATPASE (EC 3.6.1.5).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SPRAGUE DAWLEY; TISSUE-BRAIN;
 RX MEDLINE; 98031057.
 RA KEGEL B., BRAUN N., HEINE P., MALISZEWSKI C.R., ZIMMERMANN H.;
 RT "An ecto-ATPase and an ecto-ATP diphosphohydrolase are expressed in
 rat brain.";
 RL Neuroparmacology 36:1189-1200(1997).
 CC -1- FUNCTION: HYDROLYSES EXTRACELLULAR ATP TO AMP.
 CC -1- CATALYTIC ACTIVITY: ATP + 2 H₂O = AMP + 2 PHOSPHATE.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- TISSUE SPECIFICITY: HIGHEST LEVELS IN HEART, KIDNEY, AND SPLEEN.
 CC -1- STRONG SIGNALS IN THYMUS, LUNG, SKELETAL MUSCLE, AND BRAIN. VERY
 CC WEAK SIGNAL IN LIVER.
 CC -1- PMW: HAS PROBABLY VARIOUS DISULFIDE-BONDS.
 CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.
 DR EMBL; Y11835; CAI72533.1; -
 DR PROSITE; PS01238; GDAL_CD39_NTPASE; 1.
 DR PFMW; PF01150; GDAL_CD39; 1.
 KW Hydrolase; Transmembrane; Antigen; Glycoprotein.

FT DOMAIN 1 4 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 5 25 POTENTIAL.
 FT DOMAIN 22 25 POLY-LEU.
 FT DOMAIN 26 462 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 463 483 POTENTIAL.
 FT DOMAIN 467 470 POLY-LEU.
 FT DOMAIN 484 495 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 129 129 POTENTIAL.
 FT CARBOHYD 129 129 POTENTIAL.
 FT CARBOHYD 294 294 POTENTIAL.
 FT CARBOHYD 306 306 POTENTIAL.
 FT CARBOHYD 319 319 POTENTIAL.
 FT CARBOHYD 378 378 POTENTIAL.
 FT CARBOHYD 443 443 POTENTIAL.
 SQ SEQUENCE 495 AA; 54389 MW; 24F8E16C CRC32;
 Query Match 41.6%; Score 1361; DB 11; Length 495;
 Best Local Similarity 45.5%; Pred. No. 4.39e-284;
 Matches 198; Conservative 83; Mismatches 139; Indels 15; Gaps 13;
 Db 29 TODVREPPALKGIYVDAGSSHTSMFYKMPADKENDTGIVGSHSCDVGAGGSISSAND 88
 QY 1 TONKALPENVKKGIYVDAGSSHTSLYIKMPAKENDGVHVECECKYKGGISKVYOK 60
 Db 89 PSRAGOSLVCELEOALRDY-PRDRHASTPLYLGATAGMRPENTSPKATVLEAVTOTL 147
 QY 61 VNEIGIYILDCEMERA-REVIPIRSQHETPVYLGATAGMRLNMBESBELADRYLDYVERSL 119
 Db 148 TRYPPDFRGARILISGODEVFGWNTANYLLENFI-KYGV-VGRWIRPKRGLGAMDLCG 204
 QY 120 SNYPDFOGARILITGEEGAYGMITINYLKFKSQKTRMFSIYVETNNQETFGALDLCG 179
 Db 205 ASTQTFETTS-EDPQNEVHLRLYGQHYRYTHSFLCYGQDQVLQLLASALDI--HR 261
 QY 180 ASTQVTFPQONQITSPDUALQFLRGKDYNYTHSFLCYGQDQALMOKLQDVASNE 239
 Db 262 F-H-PCMPKGYSTOVLREYVOSPTMGORPAPRAGSAIVSLSGTSMALCDIVSLRFLN 319
 QY 240 ILRDPCHFGKRYKVVNSDLTKTPCTK-REFEMLT-PFOOFELQIGNTQOCHSILELFN 297
 Db 320 ISSCFSCSPGNGVFOFPVAGNFIAFSAFYTVDFLTVMGLPVTCLKOLEBATEITCNQ 379
 QY 298 TSYCPYSCAFNGIFLPLQDGFAGFSAFYFMKFLN-LTSEKVSQEKVTEKMKF-CAQ 355
 Db 380 TMELOARVPQOQTR-LADYCAVAMFIHQLSRGYPEDRSRGRVYFEKKAADYAVGML 438
 QY 356 PWEIKTSYAGVKEKYLSEYCFSGTYILSLQGYHFTADSWEHITHFGIKIGSDAGMTL 415
 Db 439 GYMLNTNLIPADLP 453
 QY 416 GYMLNTNLIPADLP 430
 RESULT 7 PRELIMINARY; PRT; 494 AA.
 ID P79784;
 AC P79784;
 DT 01-MAY-1997 (TREMBlrel. 03, Created)
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
 DE ECTO-ATPASE (EC 3.6.1.5) (ATP PYROPHOSPHOHYDROLASE).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
 OC Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=SKELETAL MUSCLE, AND GIZZARD;
 RX MEDLINE; 97150869.
 RA KIRLEY T.L.;
 RT "Complementary DNA cloning and sequencing of the chicken muscle ecto-
 RT ATPase. Homology with the lymphoid cell activation antigen CD39.";
 RL J. Biol. Chem. 272:1076-1081(1997).
 RN [2]
 SEQUENCE OF 1-12 AND 154-176.

DT	01-NOV-1999 (TReMBLrel. 12, Last annotation update)
DE	Ecto-ATPase (EC 3.6.1.3).
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC	Eukaryota; Primates; Catarrhini; Hominoidea; Homo.
RN	[1]
RP	SEQUENCE FROM N.A.
RA	MATHEO J., HARDEN T.K., BOYER J.L.;
RT	"Functional expression of a cDNA encoding a human vascular endothelial
RL	ecto-ATPase."
DR	Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
KW	EMBL: AF144748; AAD40239.1; -.
SD	Hydrolase. 495 AA; 53665 MW; 600CEB61 CRC32;
Query Match	39.5%; Score 1293; DB 4; Length 495;
Best Local Similarity	42.3%; Pred. No. 8,92e-268;
Matches 181; Conservative	98; Mismatches 136; Indels 13; Gaps 12
Db	35 PRALKYGIIVLDAGSSHTSMFTYKMPADKENDTGIVGSHSCDVPGGGISSYADNPSCAQ 94
QY	7 PENYKIGIVLDAGSSHTSLYIKKPAKENDTGIVHQVECRVKGPSIFVQKVEIGI 66
Db	95 SLVVCLEQALADYVKEKHAGTPLYLGATAGRLINTLNPEASTSVYLAAMVHTLTQYFPDF 154
QY	67 YLIDDCMRARAEVYIPRSOHOETPVYLGATAGRLIRMESEELADRVLDVVERSLSNPDPF 126
Db	155 RGARILSGOEGVGVGWTVAVYILENF-KKGWGR-WFRRPK-GTIGAMDVGASTQITF 211
QY	127 QGARITITGOEGAGVGTITNTYLGKFSQKTRMSIVYEETNNDETFCALDVGASTQITF 186
Db	212 ETTSPA-EDRASEVQALHLYGQHYRVYTHSEFLCYGRDVL-QRLTASA-LOTNCF-H-PCW 266
QY	187 VPQNOTIESPDNALQYFRLYGKDYVYTHSEFLCYGRDQALQKLAKDIQVANSNEILRDPCE 246
Db	267 PRGSTVYLLGDYIQSCSTYMAQRPNQNSSARKSVLSSSSPHLICRLVLSLFFSSCPFS 326
QY	247 HPGKKRVVNSDYIKTCTCK--REEMTLPRQOFEIIGVQOCHOSITLFLFNTSYCPYS 304
Db	327 RCPSEGVGPVPAVGNFNAFAFYVDFELRTSMGLPVATLQOLEAAVANCOTMAQLA 386
QY	305 QCAPNGITFLPPLQDGFCAFSFTYKMFLLTSE-KVSO-EKYTEMKKCAQPMETKT 362
Db	387 RVPGQAR-LADYCAGAMFVQQLSRGCGYGFDERAFGGVIFQKRAADTVAGMALGYMLNTL 445
QY	363 SYAGVKREKYLSEKCFSTYITSLDGLGYHTADSMWEIHFTIGIKQSDAGWTIGYMLNTL 422
Db	446 NLIPADPP 453
QY	423 NMIPADPP 430
RESULT	9
ID	060495
AC	060495; PRELIMINARY; PRT; 529 AA.
DT	01-AUG-1998 (TReMBLrel. 07, Created)
DT	01-AUG-1998 (TReMBLrel. 07, Last sequence update)
DT	01-MAY-1999 (TReMBLrel. 10, Last annotation update)
DE	E-TYPE ATPASE.
GN	Hb6.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC	Eutheria; Primates; Catarrhini; Hominoidea; Homo.
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=BRAIN;
RA	SMITH T.M., KIRLEY T.L.;
RL	Biochim. Biophys. Acta 0:0-0(1998).
DR	EMBL: AF034840; AAC09236.1; -.
DR	PFAM: PF01150; GDA1_CD39; 1
SD	SEQUENCE 529 AA; 59190 MW; 14BD39CD CRC32;
Query Match	39.3%; Score 1286; DB 4; Length 529;

SEQUENCE	472 AA;	51161 MW;	56924C38 CRC32;
Query Match	33.0%;	Score 1080;	DB 4; Length 472;
Best Local Similarity	43.4%;	Pred. No. 6,40e-217;	
Matches	155;	Conservative	80; Mismatches 110; Indels 12; Gaps 11.
Db	35	PALATYGIADAGSSHTSMETIKKPAKDEKNTGYIGQHSDDVYGGGSSSTADNPSGASQ	94
Qy	7	PENYKYGIVLADAGSSHTSLYIKMPAEKENDTGVHVEEGRVKGPGISKRVOKNENIGI	66
Db	95	SLVCGLEADLADVPERFAGTPLYIGATAGARKLLNTLTPKASTSYLVIAVHTTLQYEPDF	154
Qy	67	YLTCMEARREYIPPSQOETPVLIGATAGARKLLRMESEELADRYLDVVERSLSNYPDF	126
Db	155	RCARILSGOEBGVFGWYATANYLLENFI-KYGVWR-WFRPRK-CTLGAMDIGASGATITF	211
Qy	127	QCARITGOEBGAVGMITINYLCKFSQKWTWFSIVPETNNQETFGALDIGASGATVTE	186
Db	212	ETTSRA-EDRASVQALHYGCHRYTHSLCTYGRDYL-ORLLASA-LOTGCF-H-PCW	266
Qy	187	VPQNTIESPNNALOFRLYGDYNYTHSFCLCYKDKDQMLAKDKIQVASNELLRPOCF	246
Db	267	PRGFSTOVLDDVYQSPPTMAQPOPNFNSARVSLSGSDPHLCRDLYSGIFSFSSOPFS	326
Qy	247	HPGKAYVNVNPPVAGTPTCK--KREMTLLPQOFEIIGGNTQOCHOSTLEFNSTYCPYS	304
Db	327	RCSENGVQPEVPVAGNFVAFSAFEYTVDFLRTSMGLPVATLQOLLAANVNCNQTAQ	383
Qy	305	QCAFNGIFLPLPGDGFAGFASFYPMKRLNTLSE-KVSG-EKVTENMKKRCQAQWEE	359
RESULT	12	PRELIMINARY;	PRT; 483 AA.
ID	Q9X162		
AC	Q9X162		
DT	01-NOV-1999	(Tremblrel, 12, Created)	
DT	01-NOV-1999	(Tremblrel, 12, Last sequence update)	
DT	01-NOV-1999	(Tremblrel, 12, Last annotation update)	
DE	F7A19.34	PROTEIN.	
GN	F7A19.34		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;		
OC	core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;		
OC	Arabidopsis.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CV, COLUMBIA.		
RA	FEUERSTIEL N.A., PALM C.J., CONWAY A.B., CONN L., HANSEN N.F.,		
RA	ALTAFFI H., ARAUJO R., HUIZAR L., ROWLEY D., BUEHLER E., DUNN P.,		
RA	GAOZALEZ A., KREMENTSTSAIA I., KIM C., LENZ C., LI J., LIU S.,		
RA	LUNOS S., SCHWARTZ J., SHINN P., TORIOWI M., VYSTOSKAIA V.S.,		
RA	WALKER M., YU G., ECKER J., THEOLOGIS A., DAVIS R.W.,		
RL	Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AC007576; AAD39311.1;		
SO	SEQUENCE 483 AA; 53425 MW; B134313C CRC32;		
Query Match	18.0%;	Score 590;	DB 10; Length 483;
Best Local Similarity	30.9%;	Pred. No. 4.01e-102;	
Matches	131;	Conservative	104; Mismatches 154; Indels 35; Gaps 25.
Db	67	KLRYSVLIDAGSSGTRVHFGWYFESGKPVDFPEGEKHYANL-KLT-PGLSSYADNPEGAS	124
Qy	9	NVKYGVLDADAGSSHTSLYIK-W-PAEKEN-DTGVHGVVECRVKGPGISGFQKVEIG	65
Db	125	VSVYKLVFAQRIPIKRFRRSDRLMATAGMRL--EVP-VQDQILEVTRRVLRSSGEM	181
Qy	66	YLTDCMERAREVYIPRSOHOETPYLLATATAGMRLRMESEELADRYLDVVERSLSNPFD	125
Db	182	FREDMANISGSDGSIYMTANALSLG--T--D--PLET-----T-GIVEIGGSAQ	229
Qy	126	FQG--ARITITGOEBGANGMITINYLCKFSQKWTWFSIVPETNNQETFGALDIGASGATQ	183
Db	230	VTFVSSSEH-V-PPVSRITIA-VGNISYTIYSHSFLDYGKDAAL-KTLEKLONSANSNTV-	284

QY	184	VFVFQNONITESPNDALOFRLGK-DYVNVYHSFCTCYCKDDALMOKLAKKDQVANSNETLR	242
Db	285	DGVNDDPCPTPKGYIYDITMSKNYSSGFLADESKLKSLOAQNFSKRSRATPALKEGEN	344
QY	243	DCGFHDPGRKKVYVNSDYLLKTPCTKRFEMTLPRQOEIOGICIGNYQOCHSIIELFNTS--Y	300
Db	345	CLYEHCSIGSFNTPDLQGSFLATASFYTYANFEEL-EEKGMLSELIPAGKRYCGEEMSKL	403
QY	301	CYFSOCAFENGLPIPLPDQDFGAFSAFYFVMKFINLTSEKVSQEKVTEMKRKFCAQPMWEI	360
Db	404	ILIEFYTTDEBYLRGCFESAAITISMLHDSLSGLDLD-ESTIYASAGAGKHLPLDMAGAF	462
QY	361	KTSYAGVKEKYLSECFSGTYILSLLLQGYHTFADSWEHIFIKIGQSD-A-QWTLG-Y	417
Db	463	ILDV 466	
QY	418	MLNL 421	
RESULT	13	PRELIMINARY; PRT; 609 AA.	
ID	015092		
AC	015092		
DT	01-JAN-1998	(TREMBLrel. 05, Created)	
DT	01-NOV-1998	(TREMBLrel. 08, Last sequence update)	
DT	01-NOV-1998	(TREMBLrel. 08, Last annotation update)	
DE	GUANOSINE-DIPHOSPHATASE LIKE PROTEIN (K1AA0392).		
GN	K1AA0392.		
OS	Homo sapiens (Human).		
OC	Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
OC	Eutheria; Primates; Catarrhini; Hominoidea; Homo.		
NC	[1]		
RP	SEQUENCE FROM N.A.		
RP	TISSUE-BRAIN:		
RX	MEDLINE; 98225229.		
RA	WANG T.F., GUIDOTTI G.;		
RT	"Golgi localization and functional expression of human uridine		
RT	diphosphatase.";		
RL	J. Biol. Chem. 273:11392-11399(1998).		
RL	[2]		
RP	SEQUENCE OF 60-609 FROM N.A.		
RC	TISSUE-BRAIN:		
RX	MEDLINE; 97349984.		
RA	NAKASE T., ISHIKAWA K., NAKAJIMA D., OHIRA M., SBKI N., MIYAJIMA N.,		
RA	TANAKA A., KOTANI H., NOMURA N., OHARA O.;		
RT	"Prediction of the coding sequences of unidentified human genes. VII.		
RT	The complete sequences of 100 new cDNA clones from brain which can		
RT	code for large proteins in vitro.";		
RL	DNA Res. 4:141-150(1997).		
DR	EMBL; AF016032; AAC17217.1; -		
DR	EMBL; AB002390; BAA21575.1; -		
DR	PFAM; PF01150; GDA1CD39; 1.		
SO	SEQUENCE 609 AA; 69491 MW; 6A366FC5 CRC32;		
Query Match	12.1%; Score 397; DB 4; Length 609;		
Best Local Similarity	12.0%; Pred. No. 1,00e-58;		
Matches	129; Conservative 120; Mismatches 155; Indels 56; Gaps 43		
Db	87	NVNTGIYVDDCGSSGRVETVCPRRHGNPHDLDIRQNRDKNRKRVYVVKIKGISFATS	146
QY	9	NVKGIVADVADSSHSRLTYKMPAEKENDTGV--VHOV-EECR--V-K-GPGISKVFQK	60
Db	147	PEKSYDYSPLILNFAEHYVPAKHKREPLYLTCAGMHLPL-ESQQA--LLEDLDLIP	203
QY	61	VNEIGVITLDMERRRREYIPRSQOEFTPVYIGATVAGMKMLMESEELADRYLDVVERLS	120
Db	204	VHFDLFSDSHAEYVSGKQEGVYMGIVLGRFEHIEDDEAVVEVNIJGSSSESAIV	263
QY	121	NYPDF-QG-ARIITGEBGAYGMITNLYLGRFSQ---KTRAF-SI-VP-YETNN---	168
Db	264	RKFRAGILDMGCVSTQIAVEYVKTREYAKNLLAEFNLCSDVHOTEHYRVYVATFLGEGG	323
QY	169	QE-FEGGADIGASTOVWF-VPOQDTE-S--PNDALQFRLYGKD--YNYTTHFLCYGK	221

ID	AC	QX	QY	QZ	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Q23	Q24	Q25	Q26	Q27	Q28	Q29	Q30	Q31	Q32	Q33	Q34	Q35	Q36	Q37	Q38	Q39	Q40	Q41	Q42	Q43	Q44	Q45	Q46	Q47	Q48	Q49	Q50	Q51	Q52	Q53	Q54	Q55	Q56	Q57	Q58	Q59	Q60	Q61	Q62	Q63	Q64	Q65	Q66	Q67	Q68	Q69	Q70	Q71	Q72	Q73	Q74	Q75	Q76	Q77	Q78	Q79	Q80	Q81	Q82	Q83	Q84	Q85	Q86	Q87	Q88	Q89	Q90	Q91	Q92	Q93	Q94	Q95	Q96	Q97	Q98	Q99	Q100	Q101	Q102	Q103	Q104	Q105	Q106	Q107	Q108	Q109	Q110	Q111	Q112	Q113	Q114	Q115	Q116	Q117	Q118	Q119	Q120	Q121	Q122	Q123	Q124	Q125	Q126	Q127	Q128	Q129	Q130	Q131	Q132	Q133	Q134	Q135	Q136	Q137	Q138	Q139	Q140	Q141	Q142	Q143	Q144	Q145	Q146	Q147	Q148	Q149	Q150	Q151	Q152	Q153	Q154	Q155	Q156	Q157	Q158	Q159	Q160	Q161	Q162	Q163	Q164	Q165	Q166	Q167	Q168	Q169	Q170	Q171	Q172	Q173	Q174	Q175	Q176	Q177	Q178	Q179	Q180	Q181	Q182	Q183	Q184	Q185	Q186	Q187	Q188	Q189	Q190	Q191	Q192	Q193	Q194	Q195	Q196	Q197	Q198	Q199	Q200	Q201	Q202	Q203	Q204	Q205	Q206	Q207	Q208	Q209	Q210	Q211	Q212	Q213	Q214	Q215	Q216	Q217	Q218	Q219	Q220	Q221	Q222	Q223	Q224	Q225	Q226	Q227	Q228	Q229	Q230	Q231	Q232	Q233	Q234	Q235	Q236	Q237	Q238	Q239	Q240	Q241	Q242	Q243	Q244	Q245	Q246	Q247	Q248	Q249	Q250	Q251	Q252	Q253	Q254	Q255	Q256	Q257	Q258	Q259	Q260	Q261	Q262	Q263	Q264	Q265	Q266	Q267	Q268	Q269	Q270	Q271	Q272	Q273	Q274	Q275	Q276	Q277	Q278	Q279	Q280	Q281	Q282	Q283	Q284	Q285	Q286	Q287	Q288	Q289	Q290	Q291	Q292	Q293	Q294	Q295	Q296	Q297	Q298	Q299	Q300	Q301	Q302	Q303	Q304	Q305	Q306	Q307	Q308	Q309	Q310	Q311	Q312	Q313	Q314	Q315	Q316	Q317	Q318	Q319	Q320	Q321	Q322	Q323	Q324	Q325	Q326	Q327	Q328	Q329	Q330	Q331	Q332	Q333	Q334	Q335	Q336	Q337	Q338	Q339	Q340	Q341	Q342	Q343	Q344	Q345	Q346	Q347	Q348	Q349	Q350	Q351	Q352	Q353	Q354	Q355	Q356	Q357	Q358	Q359	Q360	Q361	Q362	Q363	Q364	Q365	Q366	Q367	Q368	Q369	Q370	Q371	Q372	Q373	Q374	Q375	Q376	Q377	Q378	Q379	Q380	Q381	Q382	Q383	Q384	Q385	Q386	Q387	Q388	Q389	Q390	Q391	Q392	Q393	Q394	Q395	Q396	Q397	Q398	Q399	Q400	Q401	Q402	Q403	Q404	Q405	Q406	Q407	Q408	Q409	Q410	Q411	Q412	Q413	Q414	Q415	Q416	Q417	Q418	Q419	Q420	Q421	Q422	Q423	Q424	Q425	Q426	Q427	Q428	Q429	Q430	Q431	Q432	Q433	Q434	Q435	Q436	Q437	Q438	Q439	Q440	Q441	Q442	Q443	Q444	Q445	Q446	Q447	Q448	Q449	Q450	Q451	Q452	Q453	Q454	Q455	Q456	Q457	Q458	Q459	Q460	Q461	Q462	
----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	--

DT 01-MAY-1999 (TRENBLER.10, Last annotation update)
 DE PUTATIVE NUCLEOSIDE TRIPHOSPHATASE.
 GN T17M13.14.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 OC core eudicots; Rosidae; eurosid II; Brassicales; Brassicaceae;
 OC Arabidopsis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA ROUNSELEY S.D., LIN X., KAUL S., SHEA T.P., FUJII C.Y., MASON T.M.,
 RA SHEN M., RONNING C.M., FRASER C.M., SOMERVILLE C.R., VENTER J.C.;
 RT "Arabidopsis thaliana chromosome II BAC T17M13 genomic sequence.";
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC004138; AAC32915.1; -
 DR PPRM: PF011150; GDAL_CD39; 1.
 SO SEQUENCE 516 AA; 56700 MW; 51B01A53 CRC32;

Query Match	Similarity	26.8%	Pred. No. 6.31e-19;	Length 516;
Matches 117;	Conservative 115;	Mismatches 160;	Indels 45;	Gaps 31;
Db	62	LRGSLRYSVVIDGGSTGGRHIVHGRIEESKRPVEEFGRANASLKLPHGSFAFDODGA	121	10.7% Score 352; DB 10-19;
Qy	6	LPENAKYIVLDAGSSHSLIYIKMPAEKENDIGVHNVQVEECRYK-GPGISKFYQAKNEI	64	
Db	122	SVSLTELVEFAKGRVPRKGMWIELEVRMLATAGMRLEL-PVQ--EKILGVARVLKSSGF	178	
Qy	65	GIYLLDCKMERAREVILPRSQHETEPVYLIGATAGMRLLRMESELADRLVDYVERSLSNYPF	124	
Db	179	LPEDMASVIGSDSGYAVYAVYANFALGSLG-G--D--PLKT-----T-GYELGASA	226	
Qy	125	DFQG--ARITLQGEAGAGWITTYLLKLFQSKTRMWSIYEFYENMOETFGALDGGAST	182	
Db	227	QVTFWS-SEPM-PPEFSTIS-FGNVYNTLSHSEFLPHGONAAH-DKLMSL-L-SRD--	278	
Qy	183	QVTFPPQNOTTESPDNALQFRLYRK-DYNTVTHSEFLCYGKDQALMQDLADIYVANSIEL	241	
Db	279	HNSAVEPT-RE-KI--FTDPCAPKGYINDANTQK-HLSGLA-ESRSLS--DSFOAG-	327	
Qy	242	RDPCHHPQYKRVVNVSDIKTPCT-KREMTLPRQGEIIGIGNYQOCHQSILELFTSY	300	
Db	328	GNYSOCNSAALTIIODNGRLLIITAGESF-L-FGL-GEKAWLSNMISAGERFCGSDMS	384	
Qy	301	CPYSOCAFNCGI-FLPPLQDGF-GAFSAFYFPMKFLNLTSEKVSQEKTEEMAKKFCQAPWE	358	
Db	385	KLRYKDPSELHEDDLRCFCSSAYIVSLIHDTLGIPLDD-ERIGANAGGIPLDMLGAR	443	
Qy	359	EIKSYIAVQKRYIYSEYCFSGTYLLSLLOGYHPTADSMERHIFIGIKQSSDAGWLTGYM	418	
Db	444	IQQTATETSOHAASGNL 460		
Qy	419	INTLNMLPAEQPLSTPL 435		

Search completed: Fri May 5 09:34:53 2000
Job time : 468 secs.

RESULT	15		
ID	080612	PRELIMINARY;	PRT; 516 AA.
AC	080612;		
DT	01-NOV-1998	(TREMBLrel. 08, Created)	
DT	01-NOV-1998	(TREMBLrel. 08, Last sequence update)	